(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B S

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/02637 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/21174

(22) International Filing Date: 2 July 2001 (02.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/215,495	30 June 2000 (30.06.2000)	US
60/215,483	30 June 2000 (30.06.2000)	US
60/216,117	6 July 2000 (06.07.2000)	US
60/220,589	25 July 2000 (25.07.2000)	US
60/245,294	2 November 2000 (02.11.2000)	US
60/260,729	10 January 2001 (10.01.2001)	US
60/260,851	10 January 2001 (10.01.2001)	US
60/271,673	26 February 2001 (26.02.2001)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/260,851 (CIP)
Filed on	10 January 2001 (10.01.2001)
US	60/215,495 (CIP)
Filed on	30 June 2000 (30.06.2000)
US	60/271,673 (CIP)
Filed on	26 February 2001 (26.02.2001)
US	60/215,483 (CIP)
Filed on	30 June 2000 (30.06.2000)
US	60/260,729 (CIP)
Filed on	10 January 2001 (10.01.2001)
US	60/216,117 (CIP)
Filed on	6 July 2000 (06.07.2000)
US	60/220,589 (CIP)
Filed on	25 July 2000 (25.07.2000)
US	60/245,294 (CIP)
Filed on	2 November 2000 (02.11.2000)

(71) Applicant *(for all designated States except US)*: **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SPADERNA, Steven, K.** [US/US]; 261 Deerfiled Drive, Berlin, CN 06037 (US). **PADIGARU, Muralidhara** [IN/US]; 98 Montoya Circle, Branford, CT 06405 (US). **SPYTEK, Kimberly, A.** [US/US]; 28 Court Street #1, New Haven, CT 06511 (US). **CASMAN, Stacie** [US/US]; 155 Knollwood Drive, Wallingford, CT 06492 (US). **RASTELLI, Luca** [IT/US]; 52 Pepperbush Lane, Guilford, CT 06437 (US). **MAYANKAR, Uriel** [IN/US]; 35 Averill Place, Branford, CT 06405 (US). **TCHERNEV, Velizar** [BG/US]; 45 Jefferson Road #3-12, Branford, CT 06405 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/02637 A2

(54) Title: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

(57) Abstract: Disclosed herein are nucleic acid sequences that encode G-coupled protein-receptor related polypeptides. Also disclosed are polypeptides encoded by these nucleic acid sequences, and antibodies, which immunospecifically-bind to the polypeptide, as well as derivatives, variants, mutants, or fragments of the aforementioned polypeptide, polynucleotide, or antibody. The inventio further discloses therapeutic, diagnostic and research methods for diagnosis, treatment, and prevention of disorders involving any one of these novel human nucleic acids and proteins.

NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

BACKGROUND

The invention generally relates to novel GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8, GPCR9 and GPCR10 nucleic acids and polypeptides encoded therefrom. More specifically, the invention relates to nucleic acids encoding novel polypeptides, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

BACKGROUND OF THE INVENTION

The invention generally relates to nucleic acids and polypeptides. More particularly, the invention relates to nucleic acids encoding novel G-protein coupled receptor (GPCR) polypeptides, as well as vectors, host cells, antibodies, and recombinant methods for producing these nucleic acids and polypeptides.

SUMMARY OF THE INVENTION

The invention is based in part upon the discovery of nucleic acid sequences encoding novel polypeptides. The novel nucleic acids and polypeptides are referred to herein as GPCRX, or GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8, GPCR9 and GPCR10 nucleic acids and polypeptides. These nucleic acids and polypeptides, as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be collectively designated as "GPCRX" nucleic acid or polypeptide sequences.

In one aspect, the invention provides an isolated GPCRX nucleic acid molecule encoding a GPCRX polypeptide that includes a nucleic acid sequence that has identity to the nucleic acids disclosed in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27. In some embodiments, the GPCRX nucleic acid molecule will hybridize under stringent conditions to a nucleic acid sequence complementary to a nucleic acid molecule that includes a protein-coding sequence of a GPCRX nucleic acid sequence. The invention also includes an isolated nucleic acid that encodes a GPCRX polypeptide, or a fragment, homolog, analog or derivative thereof. For example, the nucleic acid can encode a polypeptide at least 80% identical to a polypeptide comprising the amino acid sequences of SEQ ID NOS:2, 4, 6, 8, 10,

12, 14, 16, 18, 21, 23, 25 and 28. The nucleic acid can be, for example, a genomic DNA fragment or a cDNA molecule that includes the nucleic acid sequence of any of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27.

Also included in the invention is an oligonucleotide, *e.g.*, an oligonucleotide which includes at least 6 contiguous nucleotides of a GPCRX nucleic acid (*e.g.*, SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27) or a complement of said oligonucleotide.

Also included in the invention are substantially purified GPCRX polypeptides (SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28). In certain embodiments, the GPCRX polypeptides include an amino acid sequence that is substantially identical to the amino acid sequence of a human GPCRX polypeptide.

The invention also features antibodies that immunoselectively bind to GPCRX polypeptides, or fragments, homologs, analogs or derivatives thereof.

In another aspect, the invention includes pharmaceutical compositions that include therapeutically- or prophylactically-effective amounts of a therapeutic and a pharmaceutically-acceptable carrier. The therapeutic can be, *e.g.*, a GPCRX nucleic acid, a GPCRX polypeptide, or an antibody specific for a GPCRX polypeptide. In a further aspect, the invention includes, in one or more containers, a therapeutically- or prophylactically-effective amount of this pharmaceutical composition.

In a further aspect, the invention includes a method of producing a polypeptide by culturing a cell that includes a GPCRX nucleic acid, under conditions allowing for expression of the GPCRX polypeptide encoded by the DNA. If desired, the GPCRX polypeptide can then be recovered.

In another aspect, the invention includes a method of detecting the presence of a GPCRX polypeptide in a sample. In the method, a sample is contacted with a compound that selectively binds to the polypeptide under conditions allowing for formation of a complex between the polypeptide and the compound. The complex is detected, if present, thereby identifying the GPCRX polypeptide within the sample.

The invention also includes methods to identify specific cell or tissue types based on their expression of a GPCRX.

Also included in the invention is a method of detecting the presence of a GPCRX nucleic acid molecule in a sample by contacting the sample with a GPCRX nucleic acid probe or primer, and detecting whether the nucleic acid probe or primer bound to a GPCRX nucleic acid molecule in the sample.

In a further aspect, the invention provides a method for modulating the activity of a GPCRX polypeptide by contacting a cell sample that includes the GPCRX polypeptide with a compound that binds to the GPCRX polypeptide in an amount sufficient to modulate the activity of said polypeptide. The compound can be, *e.g.*, a small molecule, such as a nucleic acid, peptide, polypeptide, peptidomimetic, carbohydrate, lipid or other organic (carbon containing) or inorganic molecule, as further described herein.

Also within the scope of the invention is the use of a therapeutic in the manufacture of a medicament for treating or preventing disorders or syndromes including, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, or other disorders related to cell signal processing and metabolic pathway modulation. The therapeutic can be, *e.g.*, a GPCRX nucleic acid, a GPCRX polypeptide, or a GPCRX-specific antibody, or biologically-active derivatives or fragments thereof.

For example, the compositions of the present invention will have efficacy for treatment of patients suffering from: developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability Disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, Retinal diseases including those involving photoreception, Cell growth rate disorders; cell shape disorders, feeding disorders; control of feeding; potential obesity due to over-eating; potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; Albright Hereditary Ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation. Dentatorubro-pallidoluysian atrophy (DRPLA) Hypophosphatemic rickets, autosomal dominant (2) Acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders of the like.

The polypeptides can be used as immunogens to produce antibodies specific for the invention, and as vaccines. They can also be used to screen for potential agonist and

antagonist compounds. For example, a cDNA encoding GPCRX may be useful in gene therapy, and GPCRX may be useful when administered to a subject in need thereof. By way of nonlimiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to Neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease; multiple sclerosis; and Treatment of Albright Hereditary Ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, allergies, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders.

The invention further includes a method for screening for a modulator of disorders or syndromes including, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders or other disorders related to cell signal processing and metabolic pathway modulation. The method includes contacting a test compound with a GPCRX polypeptide and determining if the test compound binds to said GPCRX polypeptide. Binding of the test compound to the GPCRX polypeptide indicates the test compound is a modulator of activity, or of latency or predisposition to the aforementioned disorders or syndromes.

Also within the scope of the invention is a method for screening for a modulator of activity, or of latency or predisposition to an disorders or syndromes including, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders or other disorders related to cell signal processing and metabolic pathway modulation by administering a test compound to a test animal at increased risk for the aforementioned disorders or syndromes. The test animal expresses a recombinant polypeptide encoded by a GPCRX nucleic acid. Expression or activity of GPCRX polypeptide is then measured in the test animal, as is expression or activity of the protein in a control animal which recombinantly-

expresses GPCRX polypeptide and is not at increased risk for the disorder or syndrome. Next, the expression of GPCRX polypeptide in both the test animal and the control animal is compared. A change in the activity of GPCRX polypeptide in the test animal relative to the control animal indicates the test compound is a modulator of latency of the disorder or syndrome.

In yet another aspect, the invention includes a method for determining the presence of or predisposition to a disease associated with altered levels of a GPCRX polypeptide, a GPCRX nucleic acid, or both, in a subject (*e.g.*, a human subject). The method includes measuring the amount of the GPCRX polypeptide in a test sample from the subject and comparing the amount of the polypeptide in the test sample to the amount of the GPCRX polypeptide present in a control sample. An alteration in the level of the GPCRX polypeptide in the test sample as compared to the control sample indicates the presence of or predisposition to a disease in the subject. Preferably, the predisposition includes, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders. Also, the expression levels of the new polypeptides of the invention can be used in a method to screen for various cancers as well as to determine the stage of cancers.

In a further aspect, the invention includes a method of treating or preventing a pathological condition associated with a disorder in a mammal by administering to the subject a GPCRX polypeptide, a GPCRX nucleic acid, or a GPCRX-specific antibody to a subject (*e.g.*, a human subject), in an amount sufficient to alleviate or prevent the pathological condition. In preferred embodiments, the disorder, includes, *e.g.*, diabetes, metabolic disturbances associated with obesity, the metabolic syndrome X, anorexia, wasting disorders associated with chronic diseases, metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders.

In yet another aspect, the invention can be used in a method to identity the cellular receptors and downstream effectors of the invention by any one of a number of techniques commonly employed in the art. These include but are not limited to the two-hybrid system, affinity purification, co-precipitation with antibodies or other specific-interacting molecules.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention

belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention is based, in part, upon the discovery of novel nucleic acid sequences that encode novel polypeptides. The novel nucleic acids and their encoded polypeptides are referred to individually as GPCR1, GPCR2, GPCR3, GPCR4, GPCR5, GPCR6, GPCR7, GPCR8, GPCR9 and GPCR10. The nucleic acids, and their encoded polypeptides, are collectively designated herein as "GPCRX".

The novel GPCRX nucleic acids of the invention include the nucleic acids whose sequences are provided in Tables 1A, 1C, 2A, 2C, 3A, 4A, 5A, 5C, 6A, 6C, 7A, 8A, 9A, 9B and 10A, inclusive, or a fragment, derivative, analog or homolog thereof. The novel GPCRX proteins of the invention include the protein fragments whose sequences are provided in Tables 1B, 1D, 2B, 2D, 3B, 4B, 5B, 5D, 6B, 7B, 8B, 9B and 10B, inclusive. The individual GPCRX nucleic acids and proteins are described below. Within the scope of this invention is a method of using these nucleic acids and peptides in the treatment or prevention of a disorder related to cell signaling or metabolic pathway modulation.

The GPCRX proteins of the invention have a high homology to the 7tm_1 domain (PFam Acc. No. pfam00001). The 7tm_1 domain from the 7 transmembrane receptor family, which includes a number of different proteins, including, for example, serotonin receptors, dopamine receptors, histamine receptors, andrenergic receptors, cannabinoid receptors, angiotensin II receptors, chemokine receptors, opioid receptors, G-protein coupled receptor (GPCR) proteins, olfactory receptors (OR), and the like. Some proteins and the Protein Data Base Ids/gene indexes include, for example: rhodopsin (129209); 5-hydroxytryptamine receptors; (112821, 8488960, 112805, 231454, 1168221, 398971, 112806); G protein-coupled receptors (119130, 543823, 1730143, 132206, 137159, 6136153, 416926, 1169881, 136882, 134079); gustatory receptors (544463, 462208); c-x-c chemokine receptors (416718, 128999,

416802, 548703, 1352335); opsins (129193, 129197, 129203); and olfactory receptor-like proteins (129091, 1171893, 400672, 548417).

Because of the close homology among the members of the GPCRX family, proteins that are homologous to any one member of the family are also largely homologous to the other members, except where the sequences are different as shown below.

The similarity information for the GPCRX proteins and nucleic acids disclosed herein suggest that GPCR1-GPCR10 may have important structural and/or physiological functions characteristic of the Olfactory Receptor family and the GPCR family. Therefore, the nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications and as a research tool. These include serving as a specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene delivery/gene ablation), and (v) a composition promoting tissue regeneration in vitro and in vivo (vi) biological defense weapon.

G-Protein Coupled Receptor proteins (GPCRs) have been identified as a large family of G protein-coupled receptors in a number of species. These receptors share a seven transmembrane domain structure with many neurotransmitter and hormone receptors, and are likely to underlie the recognition and G-protein-mediated transduction of various signals. Human GPCR generally do not contain introns and belong to four different gene subfamilies, displaying great sequence variability. These genes are dominantly expressed in olfactory epithelium. See, *e.g.*, Ben-Arie et al., *Hum. Mol. Genet.* 1994 3:229-235; and, Online Mendelian Inheritance in Man (OMIM) entry # 164342 (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?).

The olfactory receptor (OR) gene family constitutes one of the largest GPCR multigene families and is distributed among many chromosomal sites in the human genome. See Rouquier et al., *Hum. Mol. Genet.* 7(9):1337-45 (1998); Malnic et al., *Cell* 96:713-23 (1999). Olfactory receptors constitute the largest family among G protein-coupled receptors, with up to 1000 members expected. See Vanderhaeghen et al., *Genomics* 39(3):239-46 (1997); Xie et al., *Mamm. Genome* 11(12):1070-78 (2000); Issel-Tarver et al., *Proc. Natl. Acad. Sci. USA* 93(20):10897-902 (1996). The recognition of odorants by olfactory receptors is the first stage in odor discrimination. See Krautwurst et al., *Cell* 95(7):917-26 (1998); Buck et al., *Cell* 65(1):175-87 (1991). Many ORs share some characteristic sequence motifs and

have a central variable region corresponding to a putative ligand binding site. See Issel-Tarver et al., *Proc. Natl. Acad. Sci. USA* 93:10897-902 (1996).

Other examples of seven membrane spanninG proteins that are related to GPCRs are chemoreceptors. See Thomas et al., Gene 178(1-2):1-5 (1996). Chemoreceptors have been identified in taste, olfactory, and male reproductive tissues. See *id.*; Walensky et al., *J. Biol. Chem.* 273(16):9378-87 (1998); Parmentier et al., *Nature* 355(6359):453-55 (1992); Asai et al., *Biochem. Biophys. Res. Commun.* 221(2):240-47 (1996).

The GPCRX nucleic acids and proteins are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further below. For example, a cDNA encoding the GPCR (or olfactory-receptor) like protein may be useful in gene therapy, and the receptor -like protein may be useful when administered to a subject in need thereof. The nucleic acids and proteins of the invention are also useful in potential therapeutic applications used in the treatment of developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, retinal diseases including those involving photoreception, cell growth rate disorders, cell shape disorders, feeding disorders, potential obesity due to over-eating, potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, allergies, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright hereditary ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, benign prostatic hypertrophy, psychotic and neurological disorders (including anxiety, schizophrenia, manic depression, delirium, dementia, and severe mental retardation), dentatorubro-pallidoluysian atrophy (DRPLA), hypophosphatemic rickets, autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders. Other GPCR-related diseases and disorders are contemplated.

The polypeptides can be used as immunogens to produce antibodies specific for the invention, and as vaccines. They can also be used to screen for potential agonist and antagonist compounds. For example, a cDNA encoding the GPCR-like protein may be useful in gene therapy, and the GPCR-like protein may be useful when administered to a subject in need thereof. By way of nonlimiting example, the compositions of the present invention will

have efficacy for treatment of patients suffering from developmental diseases, MHCII and III diseases (immune diseases), taste and scent detectability disorders, Burkitt's lymphoma, corticoneurogenic disease, signal transduction pathway disorders, retinal diseases including those involving photoreception, cell growth rate disorders, cell shape disorders, feeding disorders, potential obesity due to over-eating, potential disorders due to starvation (lack of appetite), noninsulin-dependent diabetes mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections (particularly infections caused by HIV-1 or HIV-2), pain, cancer (including but not limited to neoplasm; adenocarcinoma; lymphoma; prostate cancer; uterus cancer), anorexia, bulimia, asthma, allergies, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright hereditary ostoeodystrophy, angina pectoris, myocardial infarction, ulcers, benign prostatic hypertrophy, psychotic and neurological disorders (including anxiety, schizophrenia, manic depression, delirium, dementia, and severe mental retardation), dentatorubro-pallidoluysian atrophy (DRPLA), hypophosphatemic rickets, autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as Huntington's disease or Gilles de la Tourette syndrome and/or other pathologies and disorders. The novel nucleic acid encoding GPCR-like protein, and the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods.

GPCR1

A GPCR-like protein of the invention, referred to herein as GPCR1, is an Olfactory Receptor ("OR")-like protein. The novel GPCR1 nucleic acid sequences were identified on chromosome 11 as described in Example 1. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR1 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The following genomic clone was identified as having regions with high homology to the homolog. Genomic clone>acc:AP001804 HTG Homo sapiens chromosome 11 clone RP11-164A10 map 11q, WORKING DRAFT SEQUENCE, in unordered pieces - Homo sapiens, 165058 bp (DNA) was analyzed by Genscan and Grail software to identify exons and putative coding sequences.

Two alternative novel GPCR1 nucleic acids and encoded polypeptides are provided, namely GPCR1a and GPCR1b.

GPCR1a

In one embodiment, a GPCR1 variant is the novel GPCR1a (alternatively referred to herein as CG54326_02), which includes the 977 nucleotide sequence (SEQ ID NO:1) shown in Table 1A. A GPCR1a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 24-26 and ends with a TGA codon at nucleotides 957-959. The DNA sequence and protein sequence for a GPCR1a gene or one of its splice forms was obtained solely by exon linking. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 1A, and the start and stop codons are in bold letters.

Table 1A. GPCR1 Nucleotide Sequence (SEQ ID NO:1)

TTACACATAATACCTTAAAAGAC**ATG**GCTACTTCAAACCATTCTTCAGGGGCTGAGTTTATCCTGGC
AGGCTTGACACAACGCCCAGAACTTCAACTGCCACTCTTCCTCCTGTTCCTTGGAATATATGTGGTC
ACAGTGGTGGGGAACCTGGGCATGATCTTCTTAATTGCTCTCAGTTCTCAACTTTACCCTCCAGTGT
ATTATTTTCTCAGTCATTTGTCTTTCATTGATCTCTGCTACTCCTCTGTCATTACCCCTAAGATGCT
GGTGAACTTTGTTCCAGAGGAGAACATTATCTCCTTTCTGGAATGCATTACTCAACTTTATTTCTTC
CTTATTTTTGTAATTGCAGAAGGCTACCTTCTGACAGCCATGGAATGTGACCGTTATGTTGCTATCT
GTCGCCCACTGCTTTACAATATTGTCATGTCCCACAGGGTCTGTTCCATAATGATGGCTGTGGTATA
CTCACTGGGTTTTCTGTGGGCCACAGTCCATACTACCCGCATGTCAGTGTTGTCATTCTGTAGGTCT
CATACGGTCAGTCATTATTTTTGTGATATTCTCCCCTTATTGACTCTGTCTTGCTCCAGCACCCACA
TCAATGAGATTCTGCTGTTCATTATTGGAGGAGTTAATACCTTAGCAACTACACTGGCGGTCCTTAT
CTCTTATGCTTTCATTTTCTCTAGTATCCTTGGTATTCATTCCACTGAGGGGCAATCCAAAGCCTTT
GGCACTTGTAGCTCCCATCTCTTGGCTGTGGGCATCTTTTTTGGGTCTATAACATTCATGTATTTCA
AGCCCCCTTCCAGCACTACTATGGAAAAAGAGAAGGTGTCTTCTGTGTTCTACATCACAATAATCCC
CATGCTGAATCCTCTAATCTATAGCCTGAGGAACAAGGATGTGAAAAATGCACTGAAGAAGATGACT
AGGGGAAGGCAGTCATCC**TGA**CAAAGAGGGTTCTCATTG

The cDNA coding for the GPCR1a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof. The DNA sequence and protein sequence for a novel GPCR1 gene were obtained by exon linking and are reported here as GPCR1a.These primers and methods used to amplify GPCR1a cDNA are described in the Examples.

The GPCR1a polypeptide (SEQ ID NO:2) encoded by SEQ ID NO:1 is 311 aa in length, has a molecular weight of 34795.35 Daltons, and is presented using the one-letter amino acid code in Table 1B. The Psort profile for both GPCR1a and GPCR1b predicts that

these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR1 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P predicts a likely cleavage site for a GPCR1 peptide is between positions 41 and 42, *i.e.*, at the dash in the sequence VVG-NL.

Table 1B. GPCR1a protein sequence (SEQ ID NO:2)

```
MATSNHSSGAEFILAGLTQRPELQLPLFLLFLGIYVVTVVGNLGMIFLIALSSQLYPPVYYFLSHLS
FIDLCYSSVITPKMLVNFVPEENIISFLECITQLYFFLIFVIAEGYLLTAMECDRYVAICRPLLYNI
VMSHRVCSIMMAVVYSLGFLWATVHTTRMSVLSFCRSHTVSHYFCDILPLLTLSCSSTHINEILLFI
IGGVNTLATTLAVLISYAFIFSSILGIHSTEGQSKAFGTCSSHLLAVGIFFGSITFMYFKPPSSTTM
EKEKVSSVFYITIIPMLNPLIYSLRNKDVKNALKKMTRGRQSS
```

GPCR1b

In an alternative embodiment, a GPCR1 variant is the novel GPCR1b (alternatively referred to herein as AP001804_A), which includes the 936 nucleotide sequence (SEQ ID NO:3) shown in Table 1C. The GPCR1b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 1-3 and ends with a TGA codon at nucleotides 934-936, which are in bold letters in Table 1C.

Table 1C. GPCR1b Nucleotide Sequence (SEQ ID NO:3)

```
ATGGCTACTTCAAACCATTCTTCAGGGGCTGAGTTTATCCTGGCAGGCTTGACACAACGCCCAGAAC
TTCAACTGCCACTCTTCCTCCTGTTCCTTGGAATATATGTGGTCACAGTGGTGGGGAACCTGGGCAT
GATCTTCTTAATTGCTCTCAGTTCTCAACTTTACCCTCCAGTGTATTATTTTCTCAGTCATTTGTCT
TTCATTGATCTCTGCTACTCCTCTGTCATTACCCCTAAGATGCTGGTGAACTTTGTTCCAGAGGAGA
ACATTATCTCCTTTCTGGAATGCATTACTCAACTTTATTTCTTCCTTATTTTTGTAATTGCAGAAGG
CTACCTTCTGACAGCCATGGAATATGACCGTTATGTTGCTATCTGTCGCCCACTGCTTTACAATATT
GTCATGTCCCACAGGGTCTGTTCCATAATGATGGCTGTGGTATACTCACTGGGTTTTCTGTGGGCCA
CAGTCCATACTACCCGCATGTCAGTGTTGTCATTCTGTAGGTCTCATACGGTCAGTCATTATTTTTG
TGATATTCTCCCCTTATTGACTCTGTCTTGCTCCAGCACCCACATCAATGAGATTCTGCTGTTCATT
ATTGGAGGAGTTAATACCTTAGCAACTACACTGGCGGTCCTTATCTCTTATGCTTTCATTTTCTCTA
GTATCCTTGGTATTCATTCCACTGAGGGGCAATCCAAAGCCTTTGGCACTTGTAGCTCCCATCTCTT
GGCTGTGGGCATCTTTTTTGGGTCTATAACATTCATGTATTTCAAGCCCCCTTCCAGCACTACTATG
GAAAAAGAGAAGGTGTCTTCTGTGTTCTACATCACAATAATCCCCATGCTGAATCCTCTAATCTATA
GCCTGAGGAACAAGGATGTGAAAAATGCACTGAAGAAGATGACTAGGGGAAGGCAGTCATCCTGA
```

The GPCR1b protein (SEQ ID NO:4) encoded by SEQ ID NO:3 is 311 amino acid in length, has a molecular weight of 34855.38 Daltons, and is presented using the one-letter code in Table 1D. As with GPCR1a, the most likely cleavage site for a GPCR1b peptide is between amino acids 41 and 42, *i.e.*, at the dash in the sequence VVG-NL, based on the SignalP result.

Table 1D. GPCR1b protein sequence (SEQ ID NO:4)

MATSNHSSGAEFILAGLTQRPELQLPLFLLFLGIYVVTVVGNLGMIFLIALSSQLYPPVYYFLSHLS
FIDLCYSSVITPKMLVNFVPEENIISFLECITQLYFFLIFVIAEGYLLTAMEYDRYVAICRPLLYNI
VMSHRVCSIMMAVVYSLGFLWATVHTTRMSVLSFCRSHTVSHYFCDILPLLTLSCSSTHINEILLFI
IGGVNTLATTLAVLISYAFIFSSILGIHSTEGQSKAFGTCSSHLLAVGIFFGSITFMYFKPPSSTTM
EKEKVSSVFYITIIPMLNPLIYSLRNKDVKNALKKMTRGRQSS

GPCR1 Clones

Unless specifically addressed as GPCR1a or GPCR1b, any reference to GPCR1 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant. For example, the GPCR1 nucleic acid sequences differ at the following position: G382A. The GPCR1 polypeptides differ only at one residue, namely C120Y. The homologies shown above are shared by GPCR1b insofar as GPCR1a and 1b are homologous as shown in Table 1E and Table 1G. GPCRX residues in all following sequence alignments that differ between the individual GPCRX variants are marked with the (o) symbol above the variant residue in all alignments herein.

The amino acid sequence of GPCR1 had high homology to other proteins as shown in Table 1E.

Table 1E. BLASTX results for GPCR1

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prob P(N)
>patp:AAY90875 Human G protein-coupled receptor GTAR11-1	+1	1092	9.6e-110
>patp:AAY90877 Human G protein-coupled receptor GTAR11-3	+1	979	9.0e-98

In a search of sequence databases, it was found, for example, that the GPCR1 nucleic acid sequence has 657 of 932 bases (70%) identical to a gb:GENBANK-ID:RNOLP4 |acc:X80671.1 mRNA from *Rattus norvegicus* (*R. norvegicus* olp4 mRNA). The full GPCR1 amino acid sequence was found to have 209 of 305 amino acid residues (68%) identical to, and 253 of 305 amino acid residues (82%) similar to, the 309 amino acid residue ptnr:SPTREMBL-ACC:Q63395 protein from Rattus norvegicus (Rat) (OLFACTORY RECEPTOR). In all BLAST herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. Additional BLAST results are shown in Table 1F.

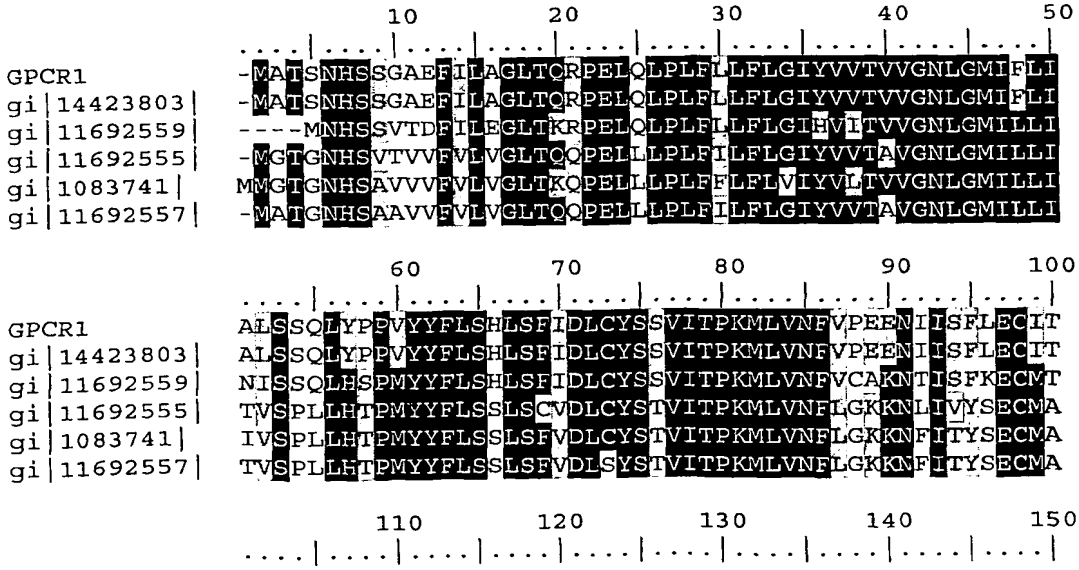
Table 1F. BLAST results for GPCR1

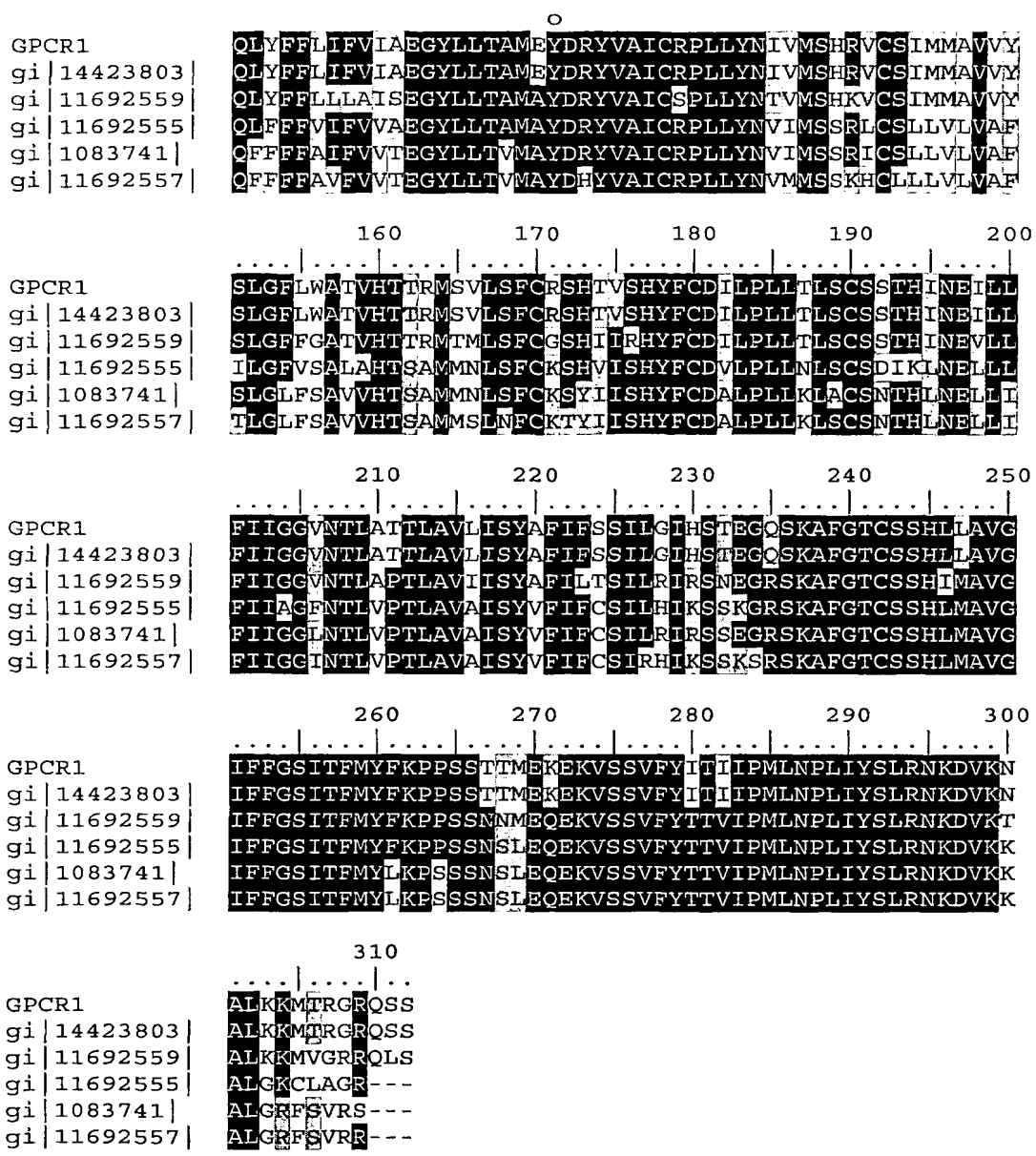
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|14423803|sp|Q9GZM6; gi|12002782|gb|AAG43386.1| (AF162668); gi|12002784|gb|AAG43387.1|AF162669_1 (AF162669)	olfactory receptor-like protein JCG2 [Homo sapiens]	311	301/311 (96%)	301/311 (96%)	e-145
gi|11692559|gb|AAG39876.1|AF282291_1 (AF282291)	gi|11692559|gb|AAG39876.1|AF282291_1 (AF282291)	308	241/307 (78%)	264/307 (85%)	e-117
gi|11692555|gb|AAG39874.1|AF282289_1 (AF282289)	odorant receptor K40 [Mus musculus]	308	209/308 (67%)	247/308 (79%)	e-101
gi|1083741|pir||S51356; gi|517366|emb|CAA56697.1| (X80671)	olfactory receptor [Rattus norvegicus]	309	202/305 (66%)	246/305 (80%)	8e-98
gi|11692557|gb|AAG39875.1|AF282290_1 (AF282290)	odorant receptor K41 [Mus musculus]	308	195/305 (63%)	241/305 (78%)	5e-92

A multiple sequence alignment is given in Table 1G, with the GPCR1 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR1 with related protein sequences disclosed in Table 1F. The residue that differs between GPCR1a and GPCR1b is marked with the (o) symbol.

Table 1G. Information for the ClustalW proteins:

1. >GPCR1; SEQ ID NO:4
2. >gi|14423803|sp|Q9GZM6|O8D2_Human Olfactory Receptor 8D2 (OR-Like Protein JCG2); SEQ ID NO:29
3. >gi|11692559|gb|AAG39876.1|AF282291_1 odorant receptor K42 [Mus musculus]; SEQ ID NO:30
4. >gi|11692555|gb|AAG39874.1|AF282289_1 odorant receptor K40 [Mus musculus]; SEQ ID NO:31
5. >gi|1083741|pir||S51356 olfactory receptor - rat; SEQ ID NO:32
6. >gi|11692557|gb|AAG39875.1|AF282290_1 odorant receptor K41 [Mus musculus]; SEQ ID NO:33





The presence of identifiable domains in the protein disclosed herein was determined by searches using algorithms such as PROSITE, Blocks, Pfam, ProDomain, Prints and then determining the Interpro number by crossing the domain match (or numbers) using the Interpro website (http:www.ebi.ac.uk/interpro/). The results indicate that the GPCR1 protein contains the following protein domain (as defined by Interpro): domain name 7tm_1 7 transmembrane receptor (rhodopsin family). DOMAIN results for GPCR1 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections.

As discussed below, all GPCRX proteins of the invention contain significant homology to the 7tm_1 domain. This indicates that the GPCRX sequence has properties similar to those of other proteins known to contain this 7tm_1 domain and similar to the properties of these domains. The 254 amino acid domain termed 7tm_1 (SEQ ID NO:34), a seven transmembrane receptor (rhodopsin family), is shown in Table 1H.

Table 1H. 7tm_1, 7 transmembrane receptor domain

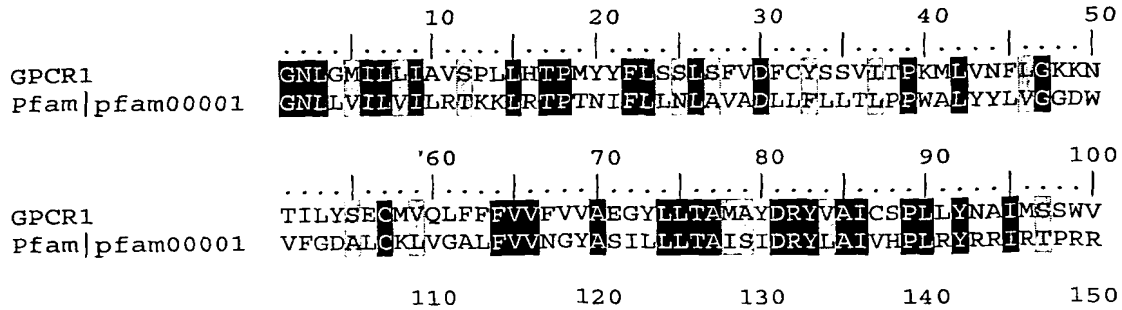
gnl|Pfam|pfam00001, 7tm_1, 7 transmembrane receptor (rhodopsin family). (SEQ ID NO:34)
GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPPWALYYLVGGDWVFGDALCKLVGALFVVNGYASILLLTAISIDRYL AIVHPLRYRRIRTPRRAKVLILLVWVLALLLSLPPLLFSWLRTVEEGNTTVCLIDFPEESVKRSYVLLSTLVGFVLPLLVILVC YTRILRTLRKRARSQRSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALLITLWLAYVNSCLNPI IY

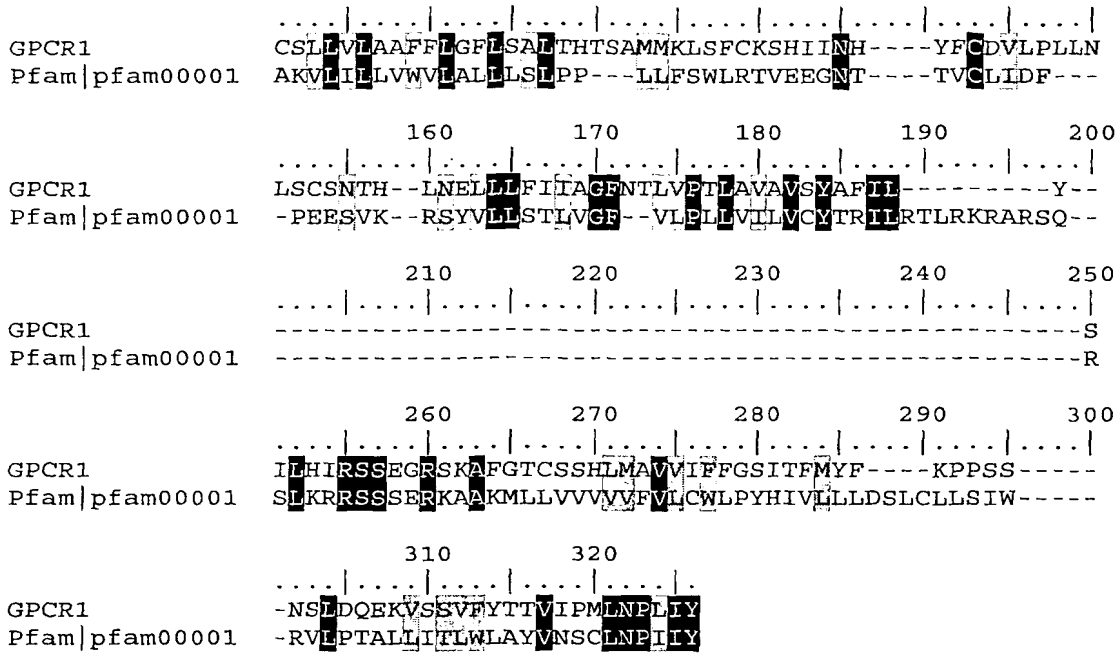
Table 1I lists the domain description from DOMAIN analysis results against GPCR1. This indicates that the GPCR1a sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34). For Table 1I and all successive DOMAIN sequence alignments, fully conserved single residues are indicated by black shading and "strong" semi-conserved residues are indicated by grey shading. The "strong" group of conserved amino acid residues may be any one of the following groups of amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW.

The DOMAIN results are listed in Table 1I with the statistics and domain description. An alignment of GPCR1 residues 41-290 (SEQ ID NOs:2 and 4) with the full 7tm_1 domain, residues 1-254 (SEQ ID NO:34), are shown in Table 1I.

Table 1I. DOMAIN results for GPCR1

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	96.3	2e-21





The nucleic acids and proteins of GPCR1 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, as described further herein.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR1 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR1 epitope is from about amino acids 10 to 20. In another embodiment, a GPCR1 epitope is from about amino acids 175 to 190. In specific embodiments, GPCR1 epitopes are from about amino acids 230 to 245, from about amino acids 258 to 273 and from about amino acids 290 to 311.

GPCR2

A second GPCR-like protein of the invention, referred to herein as GPCR2, is an Olfactory Receptor ("OR")-like protein. The novel GPCR2 nucleic acid sequences were identified on chromosome 11 as described in Example 1. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity.

Therefore it is likely that these novel GPCR2 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

The following genomic clone was identified as having regions with high homology to the homolog. Genomic clone >acc:AP001804 HTG Homo sapiens chromosome 11 clone RP11-164A10 map 11q, WORKING DRAFT SEQUENCE, in unordered pieces - Homo sapiens, 165058 bp (DNA) was analyzed by Genscan and Grail software to identify exons and putative coding sequences.

Two alternative novel GPCR2 nucleic acids and encoded polypeptides are provided, namely GPCR2a and GPCR2b.

GPCR2a

In one embodiment, a GPCR2 variant is the novel GPCR2a (alternatively referred to herein as CG54335_02), which includes the 954 nucleotide sequence (SEQ ID NO:5) shown in Table 2A. A GPCR2a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 14-16 and ends with a TGA codon at nucleotides 938-940. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 2A, and the start and stop codons are in bold letters.

Table 2A. GPCR2 Nucleotide Sequence (SEQ ID NO:5)

TGCCTAAAGAAGA**ATG**ACCATGGAAAATTATTCTATGGCAGCTCAGTTTGTCTTAGATGGTTTAACA
CAGCAAGCAGAGCTCCAGCTGCCCCTCTTCCTCCTGTTCCTGGGAATCTATGTGGTCACAGTAGTGG
GCAACCTGGGCATGATTCTCCTGATTGCAGTCAGCCCTCTACTTCACACCCCCATGTACTATTTCCT
CAGCAGCTTGTCCTTCGTCGATTTCTGCTATTCCTCTGTCATTACTCCCAAAATGCTGGTGAACTTC
CTAGGAAAGAAGAATACAATCCTTTACTCTGAGTGCATGGTCCAGCTCTTTTTCTTTGTGGTCTTTG
TGGTGGCTGAGGGTTACCTCCTGACTGCCATGGCATATGATCGCTATGTTGCCATCTGGAGCCCACT
GCTTTATAATGCGATCATGTCCTCATGGGTCTGCTCACTGCTAGTGCTGGCTGCCTTCTTCTTGGGC
TTTCTCTCTGCCTTGACTCATACAAGTGCCATGATGAAACTGTCCTTTTGCAAATCCCACATTATCA
ACCATTACTTCTGTGATGTTCTTCCCCTCCTCAATCTCTCCTGCTCCAACACACACCTCAATGAGCT
TCTACTTTTTATCATTGCGGGGTTTAACACCTTGGTGCCCACCCTAGCTGTTGCTGTCTCCTATGCC
TTCATCCTCTACAGCATCCTTCACATCCGCTCCTCAGAGGGCCGGTCCAAAGCTTTTGGAACATGCA
GCTCTCATCTCATGGCTGTGGTGATCTTCTTTGGGTCCATTACCTTCATGTATTTCAAGCCCCCTTC
AAGTAACTCCCTGGACCAGGAGAAGGTGTCCTCTGTGTTCTACACCACGGTGATCCCCATGCTGAAC
CCTTTAATATACAGTCTGAGGAATAAGGATGTGAAGAAAGCATTAAGGAAGGTCTTAGTAGGAAAAT
GAGTCCTGATTTGGGG

The sequence of GPCR2a was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. The cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The cDNA coding for the GPCR2a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. The DNA sequence and protein sequence for a novel Olfactory Receptor-like gene were obtained by exon linking and are reported here as GPCR2a.These primers and methods used to amplify GPCR2 a cDNA are described in the Examples.

The GPCR2a polypeptide (SEQ ID NO:6) encoded by SEQ ID NO:5 is 308 aa in length, has a molecular weight of 34526.32 Daltons, and is presented using the one-letter amino acid code in Table 2B. The Psort profile for both GPCR2a and GPCR2b predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR2 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P predicts a likely cleavage site for a GPCR2 peptide is between positions 41 and 42, *i.e.*, at the dash in the sequence VVG-NL.

Table 2B. GPCR2a protein sequence (SEQ ID NO:6)

```
MTMENYSMAAQFVLDGLTQQAELQLPLFLLFLGIYVVTVVGNLGMILLIAVSPLLHTPMYY
FLSSLSFVDFCYSSVITPKMLVNFLGKKNTILYSECMVQLFFFVVFVVAEGYLLTAMAYDR
YVAIWSPLLYNAIMSSWVCSLLVLAAFFLGFLSALTHTSAMMKLSFCKSHIINHYFCDVLP
LLNLSCSNTHLNELLLFIIAGFNTLVPTLAVAVSYAFILYSILHIRSSEGRSKAFGTCSSH
LMAVVIFFGSITFMYFKPPSSNSLDQEKVSSVFYTTVIPMLNPLIYSLRNKDVKKALRKVL
VGK
```

GPCR2b

In an alternative embodiment, a GPCR2 variant is the novel GPCR2b (alternatively referred to herein as AP001804_B), which includes the 927 nucleotide sequence (SEQ ID NO:7) shown in Table 2C. The GPCR2b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 1-3 and ends with a TGA codon at nucleotides 925-927, which are in bold letters in Table 2C.

Table 2C. GPCR2b Nucleotide Sequence (SEQ ID NO:7)

```
ATGACCATGGAAAATTATTCTATGGCAGCTCAGTTTGTCTTAGATGGTTTAACACAGCAAG
CAGAGCTCCAGCTGCCCCTCTTCCTCCTGTTCCTGGGAATCTATGTGGTCACAGTAGTGGG
CAACCTGGGCATGATTCTCCTGATTGCAGTCAGCCCTCTACTTCACACCCCCATGTACTAT
TTCCTCAGCAGCTTGTCCTTCGTCGATTTCTGCTATTCCTCTGTCATTACTCCCAAAATGC
TGGTGAACTTCCTAGGAAAGAAGAATACAATCCTTTACTCTGAGTGCATGGTCCAGCTCTT
TTTCTTTGTGGTCTTTGTGGTGGCTGAGGGTTACCTCCTGACTGCCATGGCATATGATCGC
TATGTTGCCATCTGTAGCCCACTGCTTTATAATGCGATCATGTCCTCATGGGTCTGCTCAC
TGCTAGTGCTGGCTGCCTTCTTCTTGGGCTTTCTCTCTGCCTTGACTCATACAAGTGCCAT
```

GATGAAACTGTCCTTTTGCAAATCCCACATTATCAACCATTACTTCTGTGATGTTCTTCCC
CTCCTCAATCTCTCCTGCTCCAACACACACCTCAATGAGCTTCTACTTTTTATCATTGCGG
GGTTTAACACCTTGGTGCCCACCCTAGCTGTTGCTGTCTCCTATGCCTTCATCCTCTACAG
CATCCTTCACATCCGCTCCTCAGAGGGCCGCTCCAAAGCTTTTGGAACATGCAGCTCTCAT
CTCATGGCTGTGGTGATCTTCTTTGGGTCCATTACCTTCATGTATTTCAAGCCCCCTTCAA
GTAACTCCCTGGACCAGGAGAAGGTGTCCTCTGTGTTCTACACCACGGTGATCCCCATGCT
GAACCCTTTAATATACAGTCTGAGGAATAAGGATGTGAAGAAAGCATTAAGGAAGGTCTTA
GTAGGAAAA**TGA**

The GPCR2b protein (SEQ ID NO:8) encoded by SEQ ID NO:7 is 308 amino acid in length, has a molecular weight of 34443.26 Daltons, and is presented using the one-letter code in Table 2D. As with GPCR2a, the most likely cleavage site for a GPCR2b peptide is between amino acids 41 and 42, *i.e.*, at the dash in the sequence VVG-NL, based on the SignalP result.

Table 2D. GPCR2b protein sequence (SEQ ID NO:8)

MTMENYSMAAQFVLDGLTQQAELQLPLFLLFLGIYVVTVVGNLGMILLIAVSPLLHTPMYY
FLSSLSFVDFCYSSVITPKMLVNFLGKKNTILYSECMVQLFFFVVFVVAEGYLLTAMAYDR
YVAICSPLLYNAIMSSWVCSLLVLAAFFLGFLSALTHTSAMMKLSFCKSHIINHYFCDVLP
LLNLSCSNTHLNELLLFIIAGFNTLVPTLAVAVSYAFILYSILHIRSSEGRSKAFGTCSSH
LMAVVIFFGSITFMYFKPPSSNSLDQEKVSSVFYTTVIPMLNPLIYSLRNKDVKKALRKVL
VGK

GPCR2 Clones

Unless specifically addressed as GPCR2a or GPCR2b, any reference to GPCR2 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant. For example, the GPCR2 nucleic acid sequences differ at the following position: G394T. The GPCR2 polypeptides differ only at one residue, namely W127C.

The amino acid sequence of GPCR2 had high homology to other proteins as shown in Table 2E.

Table 2E. BLASTX results for GPCR2

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prob P(N)
>patp:AAY90875 Human G protein-coupled receptor GTAR11-1	+1	1484	2.8e-151
>patp:AAY90879 Human G protein-coupled receptor GTAR11-1	+1	1260	1.5e-127

In a search of sequence databases, it was found, for example, that the GPCR2 nucleic acid sequence of this invention has 770 of 922 bases (83%) identical to a *Rattus norvegicus* Olfactory Receptor-like protein mRNA (GENBANK-ID:RNOLP4|acc:X80671). The full amino acid sequence of the protein of the invention was found to have 247 of 302 amino acid

residues (81%) identical to, and 261 of 302 residues (86%) similar to, the 309 amino acid residue Olfactory Receptor-like protein from *Rattus norvegicus* (SPTREMBL-ACC:Q63395). In all BLAST herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. In addition, it was found, for example, that the GPCR2 nucleic acid sequence has 778 of 927 bases (83%) identical to a gb:GENBANK-ID:AF282289| acc:AF282289.1 mRNA from Mus musculus (Mus musculus odorant receptor K40 gene, complete cds). The full amino acid sequence of the protein of the invention was found to have 259 of 304 amino acid residues (85%) identical to, and 275 of 304 amino acid residues (90%) similar to, the 308 amino acid residue ptnr:TREMBLNEW-ACC:AAG39874 protein from Mus musculus (Mouse) (ODORANT RECEPTOR K40). Additional BLAST results are shown in Table 2F.

Table 2F. BLAST results for GPCR2

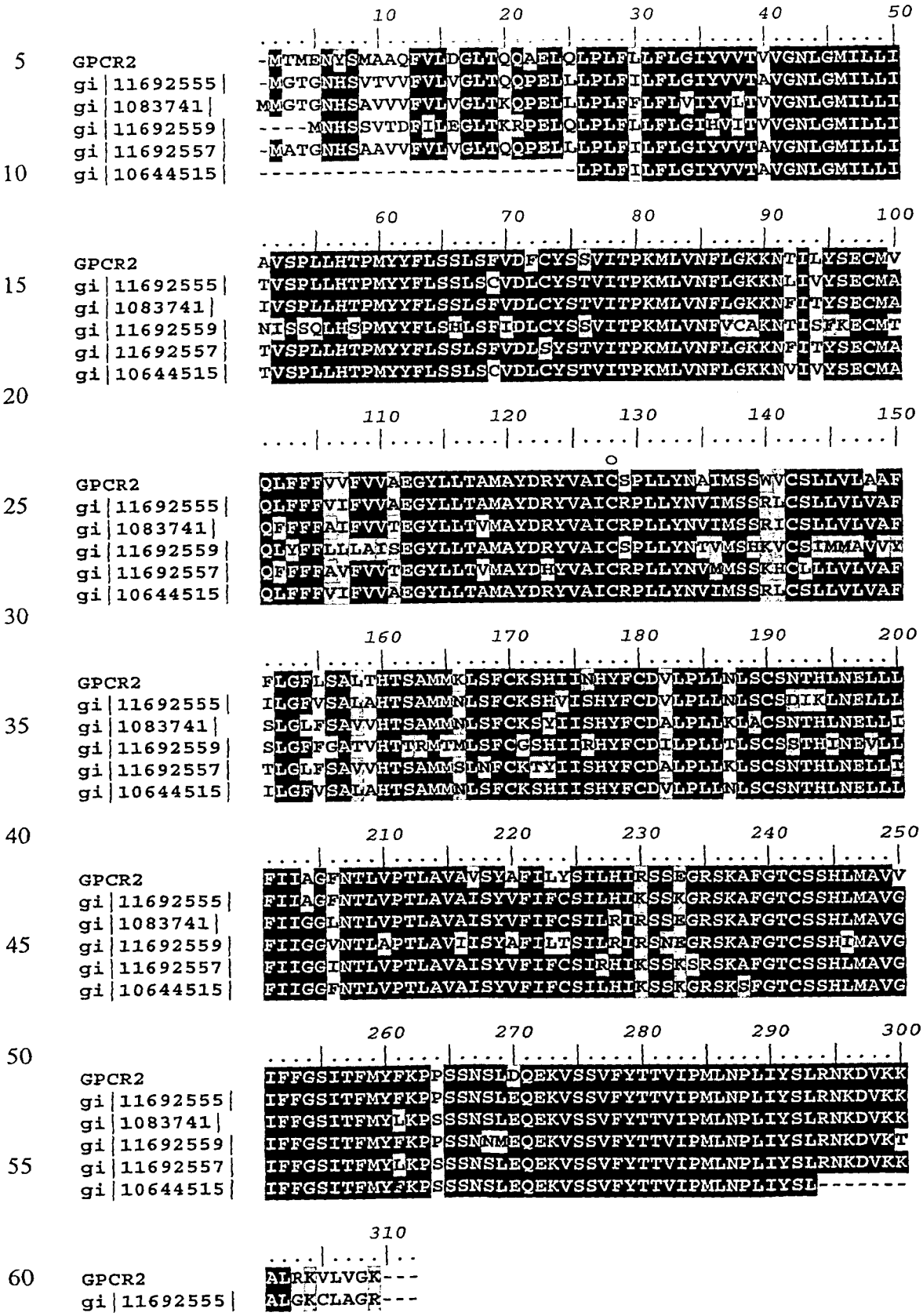
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|11692555|gb|AAG39874.1|AF282289_1 (AF282289)	odorant receptor K40 [Mus musculus]	308	233/304 (76%)	246/304 (80%)	e-109
gi|1083741|pir||S51356 gi|517366|emb|CAA56697.1| (X80671)	olfactory receptor [Rattus norvegicus]	309	223/302 (73%)	235/302 (76%)	e-102
gi|11692559|gb|AAG39876.1|AF282291_1 (AF282291)	odorant receptor K42 [Mus musculus]	308	207/301 (68%),	235/301 (77%)	e-100
gi|11692557|gb|AAG39875.1|AF282290_1 (AF282290)	odorant receptor K41 [Mus musculus]	308	219/306 (71%)	232/306 (75%)	e-100
gi|10644515|gb|AAG21322.1|AF271049_1 (AF271049)	odorant receptor [Mus musculus]	268	211/260 (81%)	220/260 (84%)	2e-95

A multiple sequence alignment is given in Table 2G, with the GPCR2 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR2 with related protein sequences of Table 2F. The residue that differs between GPCR2a and GPCR2b is marked with the (o) symbol.

Table 2G. Information for the ClustalW proteins:

1. >GPCR2; SEQ ID NO:5
2. >gi|11692555|gb|AAG39874.1|AF282289_1 odorant receptor K40 [Mus musculus]; SEQ ID NO:35
3. >gi|1083741|pir||S51356 olfactory receptor - rat; SEQ ID NO:36
4. >gi|11692559|gb|AAG39876.1|AF282291_1 odorant receptor K42 [Mus musculus]; SEQ ID NO:37
5. >gi|11692557|gb|AAG39875.1|AF282290_1 odorant receptor K41 [Mus musculus]; SEQ ID NO:38

6. >gi|10644515|gb|AAG21322.1|AF271049_1 odorant receptor [Mus musculus]; SEQ ID NO:39



```
gi|1083741|    ALGRFSVRS---
gi|11692559|   ALKKMVGRRQLS
gi|11692557|   ALGRFSVRR---
gi|10644515|   ------------
```

DOMAIN results for GPCR2 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 2H with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR2. An alignment of GPCR2 residues 41-290 (SEQ ID NO:6) with 7tm_1 residues 1-254 (SEQ ID NO:34) are shown in Table 2H.

Table 2H. DOMAIN results for GPCR2

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	85.5	4e-18

```
                        10        20        30        40        50
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           GNLGMILLIAVSPLLHTPMYYFLSSLSFVDFCYSSVITPKMLVNFLGKKN
Pfam|pfam00001  GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPFWALYYLVGGDW

                        60        70        80        90        100
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           TILYSECMVQLFFFVVFVVAEGYLLTAMAYDRYVAICSPLLYNAIMSSWV
Pfam|pfam00001  VFGDALCKLVGALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRR

                        110       120       130       140       150
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           CSLLVLAAFFLGFLSALTHTSAMMKLSFCKSHIINH----YFCDVLPLLN
Pfam|pfam00001  AKVLILLVWVLALLLSLPP---LLFSWLRTVEEGNT----TVCLIDF---

                        160       170       180       190       200
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           LSCSNTH--LNELLLFIIAGFNTLVPTLAVAVSYAFIL---------Y--
Pfam|pfam00001  -PEESVK--RSYVLLSTLVGF--VLPLLVILVCYTRILRTLRKRARSQ--

                        210       220       230       240       250
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           ----------------------------------------------S
Pfam|pfam00001  ----------------------------------------------R

                        260       270       280       290       300
                ....|....|....|....|....|....|....|....|....|....|
GPCR2           ILHIRSSEGRSKAFGTCSSHLMAVVIFFGSITFMYF----KPPSS-----
Pfam|pfam00001  SLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIW-----

                        310       320
                ....|....|....|....|....|.
GPCR2           -NSLDQEKVSSVFYTTVIPMLNPLIY
Pfam|pfam00001  -RVLPTALLITLWLAYVNSCLNPIIY
```

The nucleic acids and proteins of GPCR2 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR2 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR2 epitope is from about amino acids 5 to 20. In other specific embodiments, GPCR2 epitopes are from about amino acids 230 to 245, from about amino acids 260 to 275 and from about amino acids 285 to 308

GPCR3

The disclosed novel GPCR3 nucleic acid (SEQ ID NO:9) of 936 nucleotides (also referred to AP001804_C) is shown in Table 3A. The following genomic clone was identified as having regions with high homology to the GPCR3 homolog. Genomic clone >acc:AP001804 HTG Homo sapiens chromosome 11 clone RP11-164A10 map 11q, WORKING DRAFT SEQUENCE, in unordered pieces - Homo sapiens, 165058 bp (DNA) was analyzed by Genscan and Grail software to identify exons and putative coding sequences. An ORF begins with an ATG initiation codon at nucleotides 1-3 and ends with a TGA codon at nucleotides 934-936. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 3A, and the start and stop codons are in bold letters.

Table 3A. GPCR3 Nucleotide Sequence (SEQ ID NO:9)

ATGGCTGCTGAGAATTCCTCCTTCGTGACACAGTTTATCCTCGCAGGCTTAACTGACCAACCGGGAG
TCCAGATCCCCCTCTTCTTCCTGTTTCTAGGCTTCTACGTGGTCACTGTGGTGGGGAACCTGGGCTT
GATAACCCTGATAAGGCTCAACTCTCACTTGCACACCCCTATGTACTTCTTCCTCTATAACTTGTCC
TTCATAGATTTCTGCTATTCCAGTGTTATCACTCCCAAAATGCTGATGAGCTTTGTCTTAAAGAAGA
ACAGCATCTCCTACGCAGGGTGTATGACTCAGCTCTTCTTCTTTCTTTTCTTTGTTGTCTCTGAGTC
CTTCATCCTGTCAGCAATGGCGTATGACCGCTATGTGGCCATCTGTAACCCACTGTTGTACATGGTC
ACCATGTCTCCCCAGGTGTGTTTTCTCCTTTTGTTGGGTGTCTATGGGATGGGGTTTGCTGGGGCCA
TGGCCCACACAGCGTGCATGATGGGTGTGACCTTCTGTGCCAATAACCTTGTCAACCACTACATGTG
TGACATCCTTCCCCTTCTTGAGTGTGCTTGCACCAGCACCTATGTGAATGAGCTTGTAGTGTTTGTT
GTTGTGGGCATTGATATTGGTGTGCCCACAGTCACCATCTTCATTTCCTATGCTCTCATTCTCTCCA

GCATCTTCCACATTGATTCCACGGAGGGCAGGTCCAAAGCCTTCAGCACCTGCAGCTCCCACATAAT
TGCAGTTTCTCTGTTCTTTGGGTCAGGAGCATTCATGTACCTCAAACCCTTTTCTCTTTTAGCTATG
AACCAGGGCAAGGTGTCTTCCCTATTCTATACCACTGTGGTGCCCATGCTCAACCCATTAATTTATA
GCCTGAGGAATAAGGACGTCAAAGTTGCTCTAAAGAAAATCTTGAACAAAAATGCATTCTCC**TGA**

The GPCR3 protein (SEQ ID NO:10) encoded by SEQ ID NO:9 is 311 aa in length, has a molecular weight of 34480.27 Daltons, and is presented using the one-letter amino acid code in Table 3B. The Psort profile for GPCR3 predicts that these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR3 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a microbody (peroxisome) with a certainty of 0.300. The Signal P predicts a likely cleavage site for a GPCR3 peptide is between positions 41 and 42, *i.e.*, at the dash in the sequence VVG-NL.

Table 3B. Encoded GPCR3 protein sequence (SEQ ID NO:10)

MAAENSSFVTQFILAGLTDQPGVQIPLFFLFLGFYVVTVVGNLGLITLIRLNSHLHTPMYFFLYN
LSFIDFCYSSVITPKMLMSFVLKKNSISYAGCMTQLFFFLFFVVSESFILSAMAYDRYVAICNPL
LYMVTMSPQVCFLLLLGVYGMGFAGAMAHTACMMGVTFCANNLVNHYMCDILPLLECACTSTYVN
ELVVFVVVGIDIGVPTVTIFISYALILSSIFHIDSTEGRSKAFSTCSSHIIAVSLFFGSGAFMYL
KPFSLLAMNQGKVSSLFYTTVVPMLNPLIYSLRNKDVKVALKKILNKNAFS

The amino acid sequence of GPCR3 had high homology to other proteins as shown in Table 3C.

Table 3C. BLASTX results for GPCRC

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prob P(N)
>patp:AAY90877 Human G protein-coupled receptor GTAR11-3	+1	1172	3.2e-118
>patp:AAY90876 Human G protein-coupled receptor GTAR11-2	+1	1143	3.8e-115

In a search of sequence databases, it was found, for example, that the GPCR3 nucleic acid sequence has 659 of 929 bases (70%) identical to a *Rattus norvegicus* Olfactory Receptor-like protein mRNA (GENBANK-ID:AF010293|acc:AF010293). The full GPCR3 amino acid sequence was found to have 234 of 299 amino acid residues (78%) identical to, and 264 of 299 residues (88%) similar to, the 309 amino acid residue Olfactory Receptor-like protein from *Rattus norvegicus* (ptnr:PIR-ID:S29709). GPCR3 has 100% homology to OLFACTORY RECEPTOR 8B8 (OLFACTORY RECEPTOR TPCR85) (gi|14423794|sp|Q15620| O8B8_HUMAN [14423794]) disclosed Apr 21, 2001 on the

GenBank website. See, Vanderhaeghen,*et al.*, 1997 *Genomics* 39 (3), 239-246. GPCR3 also has homology to the proteins shown in the BLASTP data in Table 3D.

Table 3D. BLAST results for GPCR3

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|14423794|sp|Q15620|	O8B8_HUMAN OLFACTORY RECEPTOR 8B8 (OLFACTORY RECEPTOR TPCR85)	311	286/311 (91%)	286/311 (91%)	e-134
gi|11692535|gb|AAG39864.1|AF282279_1 (AF282279)	odorant receptor K21 [Mus musculus]	310	244/311 (78%)	265/311 (84%)	e-112
gi|11692541|gb|AAG39867.1|AF282282_1 (AF282282)	odorant receptor K23 [Mus musculus]	310	231/311 (74%)	258/311 (82%)	e-106
gi|423701|pir||S29709	olfactory receptor OR14 - rat	304	214/299 (71%)	242/299 (80%)	1e-97
gi|11692539|gb|AAG39866.1|AF282281_1 (AF282281)	odorant receptor K22 [Mus musculus]	309	220/311 (70%)	249/311 (79%)	6e-97

A multiple sequence alignment is given in Table 3E, with the GPCR3 protein being shown on line 1 in Table 3E in a ClustalW analysis, and comparing the GPCR3 protein with the related protein sequences shown in Table 3D. This BLASTP data is displayed graphically in the ClustalW in Table 3E.

Table 3E. ClustalW Analysis of GPCR3

1) GPCR3; SEQ ID NO:10
2) >gi|14423794|sp|Q15620|O8B8_ Human Olfactory Receptor 8B8 (OR TPCR85); SEQ ID NO:40
3) >gi|11692535|gb|AAG39864.1|AF282279_1 odorant receptor K21 [Mus musculus]; SEQ ID NO:41
4) >gi|11692541|gb|AAG39867.1|AF282282_1 odorant receptor K23 [Mus musculus]; SEQ ID NO:42
5) >gi|423701|pir||S29709 olfactory receptor OR14 - rat; SEQ ID NO:43
6) >gi|11692539|gb|AAG39866.1|AF282281_1 odorant receptor K22 [Mus musculus]; SEQ ID NO:44

```
                    10        20        30        40        50
                    |....|....|....|....|....|....|....|....|....|....|
GPCR3               MAAENSSFVTQFILAGLTDQPGVQIPLFFLFLGFYVVTVVGNLGLITLIR
gi|14423794|        MAAENSSFVTQFILAGLTDQPGVQIPLFFLFLGFYVVTVVGNLGLITLIR
gi|11692535|        MATENA-SVPEFILAGLTDQPGLRMPLFFLFLGFYMVTMVGNLGLITLIG
gi|11692541|        MTAKNS-SVTEFILAGLTDQPGLRMPLFFLFLGFYMVTVVGNLGLISLIG
gi|423701|          -------SVTEFILAGLTDQPGLRMPLFFLFLGFYMVTVVGNLIGIFLIG
gi|11692539|        MTAKNS-SVIEFILAGLTDQPGLRMPLFFLFLGFYMVTVVGNLGLISLIC

                    60        70        80        90        100
                    |....|....|....|....|....|....|....|....|....|....|
GPCR3               LNSHLHTPMYFFLYNLSFIDFCYSSVITPKMLMSFVLKKNSISYAGCMTQ
gi|14423794|        LNSHLHTPMYFFLYNLSFIDFCYSSVITPKMLMSFVLKKNSISYAGCMTQ
gi|11692535|        LNSHLHTPMYFFLFNLSLIDFCYSTVITPKMLVSFVSKKNIISYSGCMTQ
gi|11692541|        LNSHLHTPMYFFLFNLSLIDFCYSSTISPKMLMSFISKKNIISHPGCMAQ
gi|423701|          LNSHLHTPMYFFLFNLSVVDFCFSSTIIPKMLMSFISKKNIISHSGCMTQ
gi|11692539|        LNSHLHTPMYFFLFNLSVIDFCYSSTIIPKMLMNFISRKNIISHSGCMTQ
```

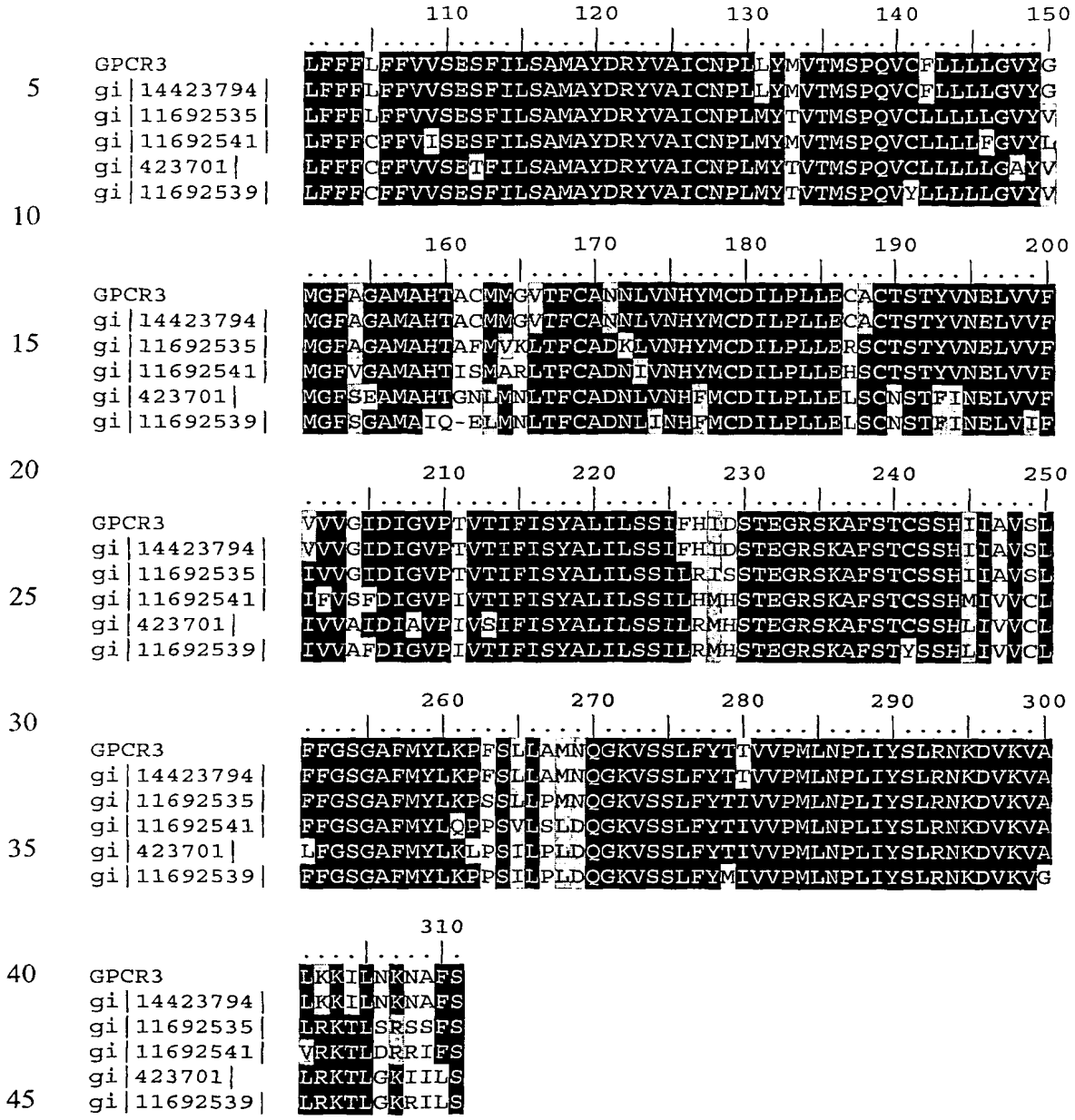
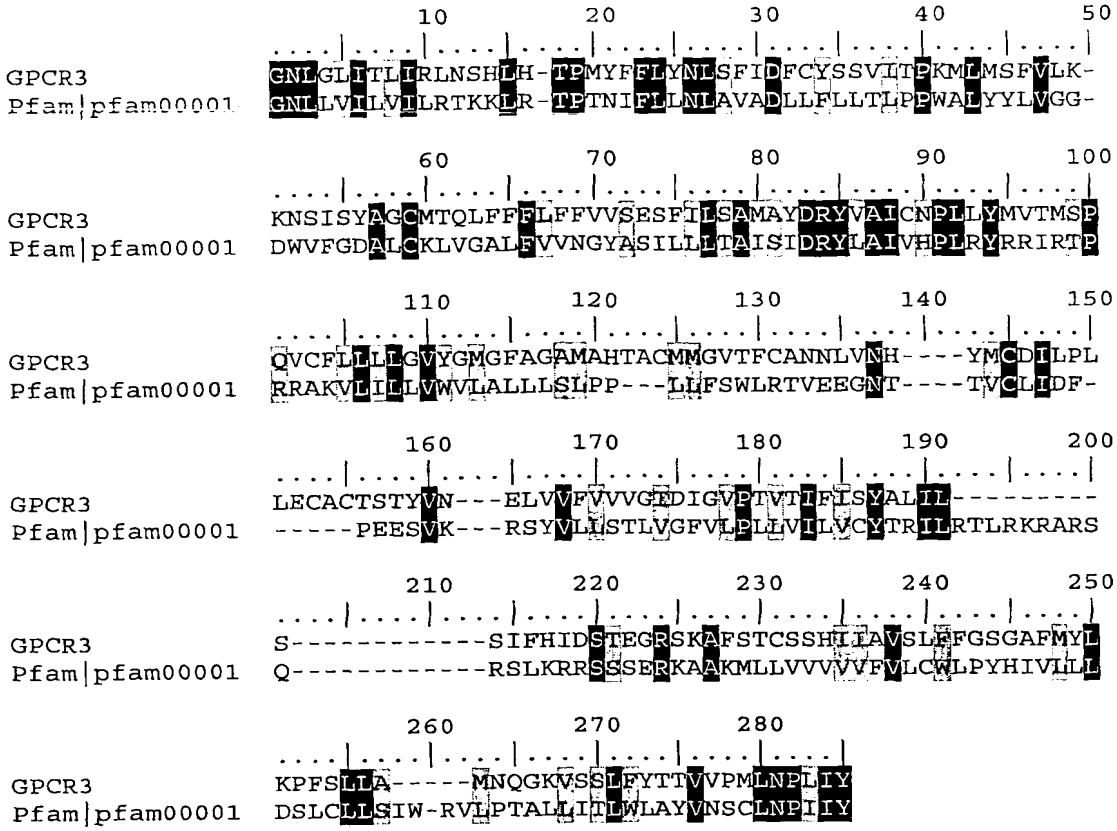


Table 3F lists the domain description from DOMAIN analysis results against GPCR3. This indicates that the GPCR3 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34) itself.

Table 3F Domain Analysis of GPCR3

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	89.0	3e-19



The nucleic acids and proteins of GPCR3 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR3 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR3 epitope is from about amino acids 10 to 20. In another embodiment, a GPCR3 epitope is from about amino acids 175 to 190. In specific embodiments, GPCR3 epitopes are from about amino acids 230 to 245and from about amino acids 285 to 311.

GPCR4

The disclosed novel GPCR4 nucleic acid (SEQ ID NO:11) of 942 nucleotides (also referred to as AP001804_D) is shown in Table 4A. The following genomic clone was identified as having regions with high homology to GPCR4: genomic clone >acc:AP001804 HTG *Homo sapiens* chromosome 11 clone RP11-164A10 map 11q, WORKING DRAFT SEQUENCE, in unordered pieces. The *Homo sapiens* 165058 bp DNA was analyzed by Genscan and Grail software programs to identify exons and putative coding sequences. A GPCR4 ORF begins with an ATG initiation codon at nucleotides 1-3 and ends with a TAA codon at nucleotides 940-942. The start and stop codons in Table 4A are in bold letters.

Table 4A. GPCR4 Nucleotide Sequence (SEQ ID NO:11)

ATGCTGGCTAGAAACAACTCCTTAGTGACTGAATTTATTCTTGCTGGATTAACAGATCATC
CAGAGTTCCAGCAACCCCTCTTTTTCCTGTTTCTAGTGGTCTACATTGTCACCATGGTAGG
CAACCTTGGCTTGATCATTCTTTTCGGTCTAAATTCTCACCTCCACACACCAATGTACTAT
TTCCTCTTCAATCTCTCCTTCATTGATCTCTGTTACTCCTCTGTTTTCACTCCCAAAATGC
TAATGAACTTTGTATCAAAAAAGAATATTATCTCCTATGTTGGGTGCATGACTCAGCTGTT
TTTCTTTCTCTTTTTTGTCATCTCTGAATGTTACATGTTGACCTCAATGGCATATGATCGC
TATGTGGCCATCTGTAATCCATTGCTGTATAAGGTCACCATGTCCCATCAGGTCTGTTCTA
TGCTCACTTTTGCTGCTTACATAATGGGATTGGCTGGAGCCACGGCCCACACCGGGTGCAT
GCTTAGACTCACCTTCTGCAGTGCTAATATCATCAACCATTACTTGTGTGACATACTCCCC
CTCCTCCAGCTTTCCTGCACCAGCACCTATGTCAACGAGGTGGTTGTTCTCATTGTTGTGG
GTATTAATATCATGGTACCCAGTTGTACCATCCTCATTTCTTATGTTTTCATTGTCACTAG
CATTCTTCATATCAAATCCACTCAAGGAAGATCAAAAGCCTTCAGTACTTGTAGCTCTCAT
GTCATTGCTCTGTCTCTGTTTTTTGGGTCAGCGGCATTCATGTATATTAAATATTCTTCTG
GATCTATGGAGCAGGGAAAAGTTTCTTCTGTTTTCTACACTAATGTGGTGCCCATGCTCAA
TCCTCTCATCTACAGTTTGAGGAACAAGGATGTCAAAGTTGCACTGAGGAAAGCTCTGATT
AAAATTCAGAGAAGAAATATATTC**TAA**

The GPCR4 protein (SEQ ID NO:12) encoded by SEQ ID NO:11 has 313 amino acid residues and is presented using the one-letter code in Table 4B. The predicted molecular weight of GPCR4 protein is approximately 35303.38 Daltons. The Psort profile for GPCR4 predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6. In alternative embodiments, GPCR4 is located in the Golgi body with a certainty of 0.4, the endoplasmic reticulum (membrane) with a certainty of 0.3 or microbodies (peroxisomes) with a certainty of 0.3. The Signal P predicts a likely cleavage site between positions 44 and 45, i.e., at the dash in the sequence NLG-LI.

Table 4B. Encoded GPCR4 protein sequence (SEQ ID NO:12)

MLARNNSLVTEFILAGLTDHPEFQQPLFFLFLVVYIVTMVGNLGLIILFGLNSHLHTPMYYFLF
NLSFIDLCYSSVFTPKMLMNFVSKKNIISYVGCMTQLFFFLFFVISECYMLTSMAYDRYVAICN

PLLYKVTMSHQVCSMLTFAAYIMGLAGATAHTGCMLRLTFCSANIINHYLCDILPLLQLSCTST
YVNEVVVLIVVGINIMVPSCTILISYVFIVTSILHIKSTQGRSKAFSTCSSHVIALSLFFGSAA
FMYIKYSSGSMEQGKVSSVFYTNVVPMLNPLIYSLRNKDVKVALRKALIKIQRRNIF

The amino acid sequence of GPCR4 had high homology to other proteins as shown in Table 4C.

Table 4C. BLASTX results for GPCR4

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prob P(N)
>patp:AAY90877 Human G protein-coupled receptor GTAR11-3	+1	1586	4.3e-162
>patp:AAY90876 Human G protein-coupled receptor GTAR11-2	+1	1544	1.2e-157

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of this invention AP001804_D has 650 of 917 bases (70%) identical to a *Rattus norvegicus* Olfactory Receptor-like protein mRNA (GENBANK-ID:AF010293|acc:AF010293). The full amino acid sequence of the protein of the invention was found to have 205 of 300 amino acid residues (68%) identical to, and 247 of 300 residues (82%) similar to, the 309 amino acid residue Olfactory Receptor-like protein from *Rattus norvegicus* (ptnr:PIR-ID:S29709). GPCR4 also has homology to the proteins shown in the BLASTP data in Table 4D.

Table 4D. BLAST results for GPCR4

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|11692545|gb|AAG39869.1|AF282284_1 (AF282284)	odorant receptor K26 [Mus musculus]	314	188/306 (61%)	220/306 (71%)	2e-87
gi|14423794|sp|Q15620|	O8B8_HUMAN OLFACTORY RECEPTOR 8B8 (OLFACTORY RECEPTOR TPCR85)	311	198/307 (64%)	230/307 (74%)	3e-86
gi|11692537|gb|AAG39865.1|AF282280_1 (AF282280)	odorant receptor K21h1 [Mus musculus]	314	189/308 (61%)	227/308 (73%)	5e-86
gi|11692541|gb|AAG39867.1|AF282282_1 (AF282282)	odorant receptor K23 [Mus musculus]	310	191/304 (62%)	223/304 (72%)	1e-85
gi|11692535|gb|AAG39864.1|AF282279_1 (AF282279)	odorant receptor K21 [Mus musculus]	310	196/304 (64%)	227/304 (74%)	2e-84

A multiple sequence alignment is given in Table 4E, with the GPCR4 protein being shown on line 1 in Table 4E in a ClustalW analysis, and comparing the GPCR4 protein with

the related protein sequences shown in Table 4D. This BLASTP data is displayed graphically in the ClustalW in Table 4E.

Table 4E. ClustalW Analysis of GPCR4

1) GPCR4; SEQ ID NO:12
2) >gi|11692545|gb|AAG39869.1|AF282284_1 odorant receptor K26 [Mus musculus]; SEQ ID NO: 45
3) >gi|14423794|sp|Q15620|O8B8_ Human Olfactory Receptor 8B8 (OR TPCR85); SEQ ID NO: 46
4) >gi|11692537|gb|AAG39865.1|AF282280_1 odorant receptor K21h1 [Mus musculus]; SEQ ID NO: 47
5) >gi|11692541|gb|AAG39867.1|AF282282_1 odorant receptor K23 [Mus musculus]; SEQ ID NO: 48
6) >gi|11692535|gb|AAG39864.1|AF282279_1 odorant receptor K21 [Mus musculus]; SEQ ID NO: 49

```
                                10        20        30        40        50
                     ....|....|....|....|....|....|....|....|....|....|....|
GPCR4                -----MLARNNSLVTEFILAGLTDHPEFQQPLFFLFLVVYIVTMVGNLGL
gi|11692545|         MMLGRMAFSNDSSVKEFILLGLTQQPELQMPLFFLFLGIYVVSMVGNLGL
gi|14423794|         -----MAAENSSFVTQFILAGLTDQPGVQIPLFFLFLGFYVVTVVGNLGL
gi|11692537|         MSQKRMAPRNSSSVTEFILVGFSNQPALQLPLFFVFLGIYVLTVIGNLGL
gi|11692541|         -----MTAKNS-SVTEFILAGLTDQPGLRMPLFFLFLGFYMVTVVGNLGL
gi|11692535|         -----MATENA-SVPEFILAGLTDQPGLRMPLFFLFLGFYMVTMVGNLGL

                               60        70        80        90       100
                     ....|....|....|....|....|....|....|....|....|....|....|
GPCR4                IILFGLNSHLHTPMYYFLFNLSFIDLCYSSVFTPKMLMNFVSKKNIISYV
gi|11692545|         IVLIVLNPHLHTPMYYFLFNLSFIDLCYSSVITPRMLVGFV-KQNIISHA
gi|14423794|         ITLIRLNSHLHTPMYFFLYNLSFIDFCYSSVITPKMLMSFVLKKNSISYA
gi|11692537|         ITLIGLNSSLHTPMYFFLFNLSFIDFCYSCVFTPKMLSDFV-SENIISYM
gi|11692541|         ISLIGLNSHLHTPMYFFLFNLSLIDFCYSSTISPKMLMSFLSKKNIISHP
gi|11692535|         ITLIGLNSHLHTPMYFFLFNLSLIDFCYSTVITPKMLVSFVSKKNIISYS

                              110       120       130       140       150
                     ....|....|....|....|....|....|....|....|....|....|....|
GPCR4                GCMTQLFFFLFFVISECYMLTSMAYDRYVAICNPLLYKVTMSHQVCSMLT
gi|11692545|         ECMTQLFFFCFFVIDECYILTAMAYDRYAAICKPLLYQVTMSHQVCHLMM
gi|14423794|         GCMTQLFFFLFFVVSESFILSAMAYDRYVAICNPLLYMVTMSPQVCFLLL
gi|11692537|         GCMTQLFFFCFFVNSECYVLVSMAYDRYVAICNPLLYTVTMSPQVCTLLM
gi|11692541|         GCMAQLFFFCFFVISESFILSAMAYDRYVAICNPLMYMVTMSPQVCLLLL
gi|11692535|         GCMTQLFFFLFFVVSESFILSAMAYDRYVAICNPLMYTVTMSPQVCLLLL

                              160       170       180       190       200
                     ....|....|....|....|....|....|....|....|....|....|....|
GPCR4                FAAYIMGLAGATAHTGCMLRLTFCSANIINHYLCDILPLLQLSCTSTYVN
gi|11692545|         VGVYVMGLVGAMAHTGSMLSLTFCDGNIINHYMCDIPPLQKLSCTSTSIN
gi|14423794|         LGVYGMGFAGAMAHTACMMGVTFCANNLVNHYMCDILPLLECACTSTYVN
gi|11692537|         FCSYVIGFAGAMAHTGSMLTLTFCDSNMIHHYLCEVLPLLQLSCTSTYAN
gi|11692541|         FGVYLMGFVGAMAHTISMARLTFCADNIVNHYMCDILPLLEHSCTSTYVN
gi|11692535|         LGVYVMGFAGAMAHTAFMVKLTFCADKLVNHYMCDILPLLERSCTSTYVN

                              210       220       230       240       250
                     ....|....|....|....|....|....|....|....|....|....|....|
GPCR4                EVVVLIVVGINIMVPSCTILISYVFIVTSILHIKSTQGRSKAFSTCSSHV
gi|11692545|         ELVVFIVVGVNVIIPSLIVFISYTLILSNILSIQSAEGRSKAFSTCGSHV
gi|14423794|         ELVVFVVVGIDIGVPTVTIFISYALILSSIFHIDSTEGRSKAFSTCSSHI
gi|11692537|         ELVFFIVVGVVITASSISIFISYALILSNILKIPSAEGRSKAFGTCGSHV
gi|11692541|         ELVVFIFVSFDIGVPLVTIFISYALILSSILHMHSTEGRSKAFSTCSSHM
gi|11692535|         ELVVFIVVGIDIGVPTVTIFISYALILSSILRISSTEGRSKAFSTCSSHI

                              260       270       280       290       300
```

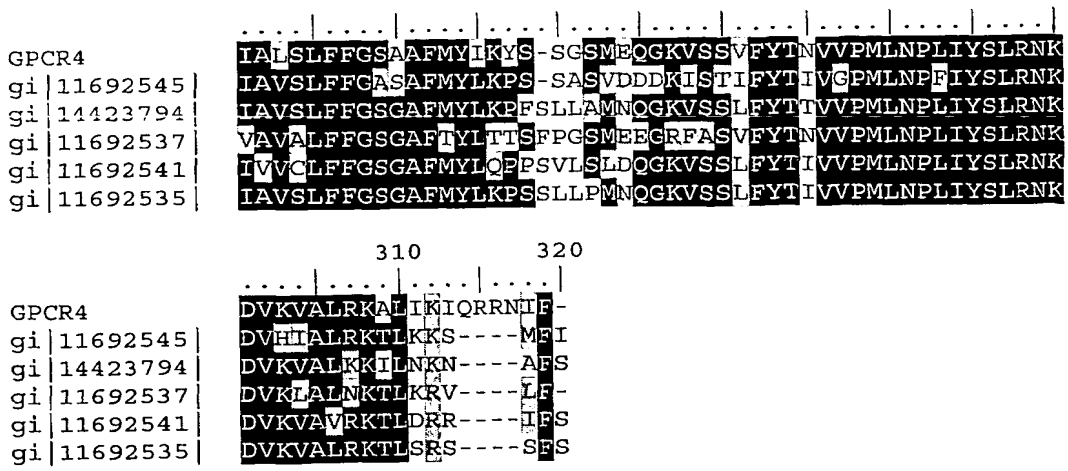
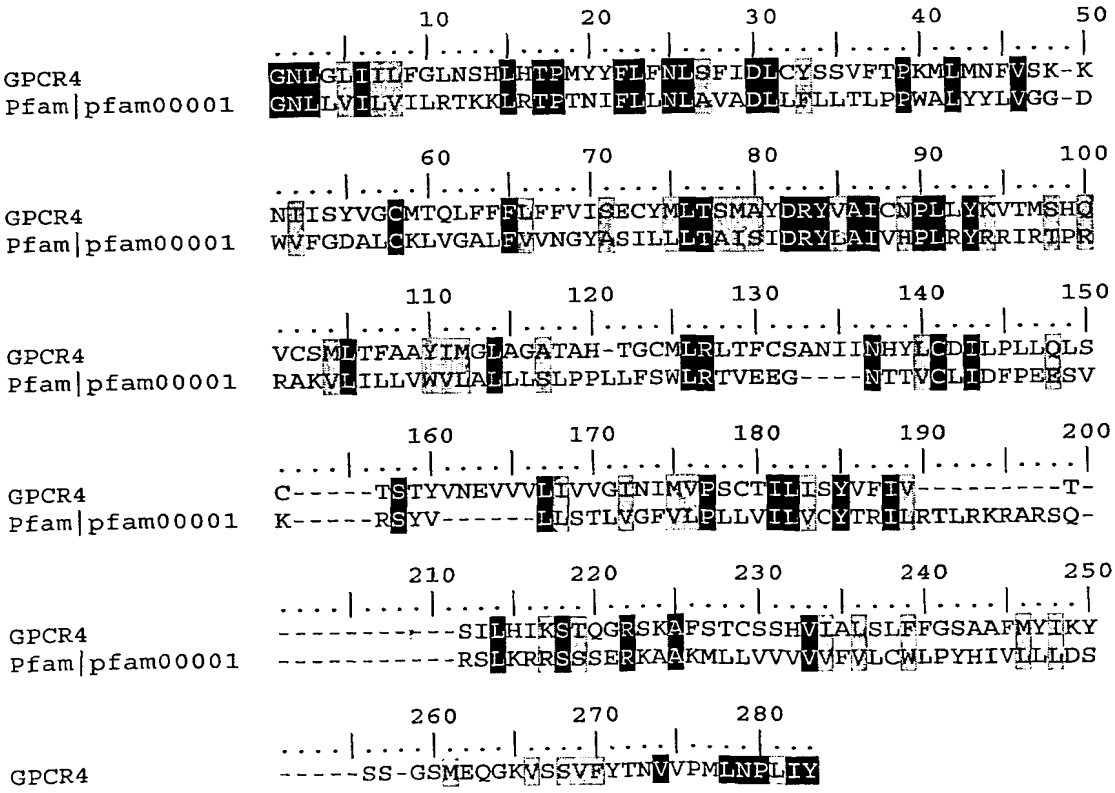


Table 4F lists the domain description from DOMAIN analysis results against GPCR4. This indicates that the GPCR4 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34) itself.

Table 4F Domain Analysis of GPCR4

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	82.8	2e-17



Pfam|pfam00001 LCLLSIW-RVLPTALLITLWLAYVNSCLNPIIY

The GPCR4 protein predicted here is similar to the "Olfactory Receptor-Like Protein Family", some members of which end up localized at the cell surface where they exhibit activity. Therefore, it is likely that this novel GPCR4 protein is available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application. The Olfactory Receptor-like GPCR4 proteins disclosed are expressed in at least the following tissues: olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types especially olfactory epithelium. Further tissue expression analysis is provided in the Examples.

The nucleic acids and proteins of GPCR4 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR4 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR4 epitope is from about amino acids 5 to 20. In additional embodiments, GPCR4 epitopes are from about amino acids 225 to 245, from about amino acids 260 to 275 and from about amino acids 290 to 313.

GPCR5

A second GPCR-like protein of the invention, referred to herein as GPCR5, is an Olfactory Receptor ("OR")-like protein. The novel GPCR5 nucleic acid sequences were identified on chromosome 11 as described in Example 1. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR5 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR5 nucleic acids and encoded polypeptides are provided, namely GPCR5a and GPCR5b.

GPCR5a

In one embodiment, a GPCR5 variant is the novel GPCR5a (alternatively referred to herein as CG56040_01), which includes the 912 nucleotide sequence (SEQ ID NO:13) shown in Table 5A. The DNA sequence and protein sequence for GPCR5a or one of its splice forms was obtained solely by exon linking. A GPCR5a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 1-3 and ends with a TGA codon at nucleotides 910-9120, shown in bold in Table 5A.

Table 5A. GPCR5 Nucleotide Sequence (SEQ ID NO:13)

ATGACTCTGAGAAACAGCTCCTCAGTGACTGAGTTTATCCTTGTGGGATTATCAGAACAGC
CAGAGCTCCAGCTCCCTCTTTTCCTTCTATTCTTAGGGATCTATGTGTTCACTGTGGTGGG
CAACTTGGGCTTGATCACCTTAATTGGGATAAATCCTAGCCTTCACACCCCCATGTACTTT
TTCCTCTTCAACTTGTCCTTTATAGATCTCTGTTATTCCTGTGTGTTTACCCCCAAAATGC
TGAATGACTTTGTTTCAGAAAGTATCATCTCTTATGTGGGATGTATGACTCAGCTATTTTT
CTTCTGTTTCTTTGTCAATTCTGAGTGCTATGTGTTGGTATCAATGGCCTATGATCGCTAT
GTGGCCATCTGCAACCCCCTGCTCTACATGGTCACCATGTCCCCAAGGGTCTGCTTTCTGC
TGATGTTTGGTTCCTATGTGGTAGGGTTTGCTGGGGCCATGGCCCACACTGGAAGCATGCT
GCGACTGACCTTCTGTGATTCCAACGTCATTGACCATTATCTGTGTGACGTTCTCCCCCTC
TTGCAGCTCTCCTGCACCAGCACCCATGTCAGTGAGCTGGTATTTTTCATTGTTGTTGGAG
TAATCACCATGCTATCCAGCATAAGCATCGTCATCTCTTACGCTTTGATACTCTCCAACAT
CCTCTGTATTCCTTCTGCAGAGGGCAGATCCAAAGCCTTTAGCACATGGGGCTCCCACATA
ATTGCTGTTGCTCTGTTTTTTGGGTCAGGGACATTCACCTACTTAACAACATCTTTTCCTG
GCTCTATGAACCATGGCAGATTTGCCTCAGTCTTTTACACCAATGTGGTTCCCATGCTTAA
CCCTTCGATCTACAGTTTGAGGAATAAGGATGATAAACTTGCCCTGGGCAAACCC**TGA**

The sequence of GPCR5a was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. The cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The cDNA coding for the GPCR5a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. The DNA sequence and protein sequence for a novel Olfactory Receptor-like gene were obtained by exon linking and are reported here as GPCR5a. These primers and methods used to amplify GPCR5 a cDNA are described in the Examples.

The GPCR5a polypeptide (SEQ ID NO:14) encoded by SEQ ID NO:13 is 303 aa in length, has a molecular weight of 33640.94 Daltons, and is presented using the one-letter amino acid code in Table 5B. The Psort profile for both GPCR5a and GPCR5b predicts that

these sequences have a signal peptide and are likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR5 polypeptide is located to the mitochondrial inner membrane with a certainty of 0.4195, the Golgi body with a certainty of 0.400, or the mitochondrial intermembrane space with a certainty of 0.3631. The Signal P predicts a likely cleavage site for a GPCR5 peptide is between positions 41 and 42, *i.e.*, at the dash in the sequence VVG-NL.

Table 5B. GPCR5a protein sequence (SEQ ID NO:14)

```
MTLRNSSSVTEFILVGLSEQPELQLPLFLLFLGIYVFTVVGNLGLITLIGINPSLHTPMYFFLFNLS
FIDLCYSCVFTPKMLNDFVSESIISYVGCMTQLFFFCFFVNSECYVLVSMAYDRYVAICNPLLYMVT
MSPRVCFLLMFGSYVVGFAGAMAHTGSMLRLTFCDSNVIDHYLCDVLPLLQLSCTSTHVSELVFFIV
VGVITMLSSISIVISYALILSNILCIPSAEGRSKAFSTWGSHIIAVALFFGSGTFTYLTTSFPGSMN
HGRFASVFYTNVVPMLNPSIYSLRNKDDKLALGKP
```

GPCR5b

In an alternative embodiment, a GPCR5 variant is the novel GPCR5b (alternatively referred to herein as AP001804_B), which includes the 930 nucleotide sequence (SEQ ID NO:15) shown in Table 5C. The GPCR5b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 1-3 and ends with a TAA codon at nucleotides 928-930, which are in bold letters in Table 5C.

Table 5C. GPCR5b Nucleotide Sequence (SEQ ID NO:15)

```
ATGACTCTGAGAAACAGCTCCTCAGTGACTGAGTTTATCCTTGTGGGATTATCAGAACAGC
CAGAGCTCCAGCTCCCTCTTTTCCTTCTATTCTTAGGGATCTATGTGTTCACTGTGGTGGG
CAACTTGGGCTTGATCACCTTAATTGGGATAAATCCTAGCCTTCACACCCCCATGTACTTT
TTCCTCTTCAACTTGTCCTTTATAGATCTCTGTTATTCCTGTGTGTTTACCCCCAAAATGC
TGAATGACTTTGTTTCAGAAAGTATCATCTCTTATGTGGGATGTATGACTCAGCTATTTTT
CTTCTGTTTCTTTGTCAATTCTGAGTGCTATGTGTTGGTATCAATGGCCTATGATCGCTAT
GTGGCCATCTGCAACCCCCTGCTCTACATGGTCACCATGTCCCCAAGGGTCTGCTTTCTGC
TGATGTTTGGTTCCTATGTGGTAGGGTTTGCTGGGGCCATGGCCCACACTGGAAGCATGCT
GCGACTGACCTTCTGTGATTCCAACGTCATTGACCATTATCTGTGTGACGTTCTCCCCCTC
TTGCAGCTCTCCTGCACCAGCACCCATGTCAGTGAGCTGGTATTTTTCATTGTTGTTGGAG
TAATCACCATGCTATCCAGCATAAGCATCGTCATCTCTTACGCTTTGATACTCTCCAACAT
CCTCTGTATTCCTTCTGCAGAGGGCAGATCCAAAGCCTTTAGCACATGGGGCTCCCACATA
ATTGCTGTTGCTCTGTTTTTTGGGTCAGGGACATTCACCTACTTAACAACATCTTTTCCTG
GCTCTATGAACCATGGCAGATTTGCCTCAGTCTTTTACACCAATGTGGTTCCCATGCTTAA
CCCTTCGATCTACAGTTTGAGGAATAAGGATGATAAACTTGCCCTGGGCAAAACCCTGAAG
AGAGTGCTCTTCTAA
```

The GPCR5b protein (SEQ ID NO:16) encoded by SEQ ID NO:15 is 309 amino acid in length, has a molecular weight of 34401.88 Daltons, and is presented using the one-letter code in Table 5D. As with GPCR5a, the most likely cleavage site for a GPCR5b peptide is

between amino acids 41 and 42, *i.e.*, at the dash in the sequence VVG-NL, based on the SignalP result.

Table 5D. GPCR5b protein sequence (SEQ ID NO:16)

```
MTLRNSSSVTEFILVGLSEQPELQLPLFLLFLGIYVFTVVGNLGLITLIGINPSLHTPMYFFLFNLS
FIDLCYSCVFTPKMLNDFVSESIISYVGCMTQLFFFCFFVNSECYVLVSMAYDRYVAICNPLLYMVT
MSPRVCFLLMFGSYVVGFAGAMAHTGSMLRLTFCDSNVIDHYLCDVLPLLQLSCTSTHVSELVFFIV
VGVITMLSSISIVISYALILSNILCIPSAEGRSKAFSTWGSHIIAVALFFGSGTFTYLTTSFPGSMN
HGRFASVFYTNVVPMLNPSIYSLRNKDDKLALGKTLKRVLF
```

GPCR5 Clones

Unless specifically addressed as GPCR5a or GPCR5b, any reference to GPCR5 is assumed to encompass all variants. The GPCR5 nucleic acid sequences have alternative 3' sequences: *i.e.*, GPCR5b has an A residue inserted at position 907 and extends 17 bp beyond the 3' end of GPCR1a. The GPCR5 polypeptides have alternative carboxyterminal sequences beginning at residue 303, wherein GPCR5 has a proline (P) as a terminal residue at position 303, and GPCR5b contains the sequence TLKRVLF at positions 303-309.

The amino acid sequence of GPCR2 had high homology to other proteins as shown in Table 5E.

Table 5E. BLASTX results for GPCR5

Sequences producing High-scoring Segment Pairs:	Reading Frame	High Score	Smallest Sum Prob P(N)
>patp:AAY90878 Human G protein-coupled receptor GTAR11-4	+1	1594	6.1e-163

The GPCR disclosed in this invention maps to chromosome 11q25. This information was assigned using OMIM, the electronic northern bioinformatics tool implemented by CuraGen Corporation, public ESTs, public literature references and/or genomic clone homologies. This was executed to derive the chromosomal mapping of the SeqCalling assemblies, Genomic clones, literature references and/or EST sequences that were included in the invention.

In a search of sequence databases, it was found, for example, that the GPCR5a nucleic acid sequence has 537 of 706 bases (76%) identical to a gb:GENBANK-ID:AF065872|acc:AF065872.1 mRNA from Homo sapiens (Homo sapiens OR8C1P pseudogene, partial sequence). The full amino acid sequence of the protein of the invention was found to have 187 of 296 amino acid residues (63%) identical to, and 233 of 296 amino acid residues (78%) similar to, the 304 amino acid residue ptnr:SPTREMBL-ACC:Q9QW36

protein from Rattus sp (OR14=ODORANT RECEPTOR In further a search of sequence databases, it was found, for example, that the GPCR5b nucleic acid sequence has 640 of 926 bases (69%) identical to a *Rattus norvegicus* Olfactory Receptor-like protein mRNA (GENBANK-ID:AF010293|acc:AF010293). The full amino acid sequence of the protein of the invention was found to have 189 of 302 amino acid residues (62%) identical to, and 238 of 302 residues (78%) similar to, the 309 amino acid residue Olfactory Receptor-like protein from *Rattus norvegicus* (ptnr:PIR-ID:S29709). Additional BLAST results are shown in Table 5F.

Table 5F. BLAST results for GPCR5

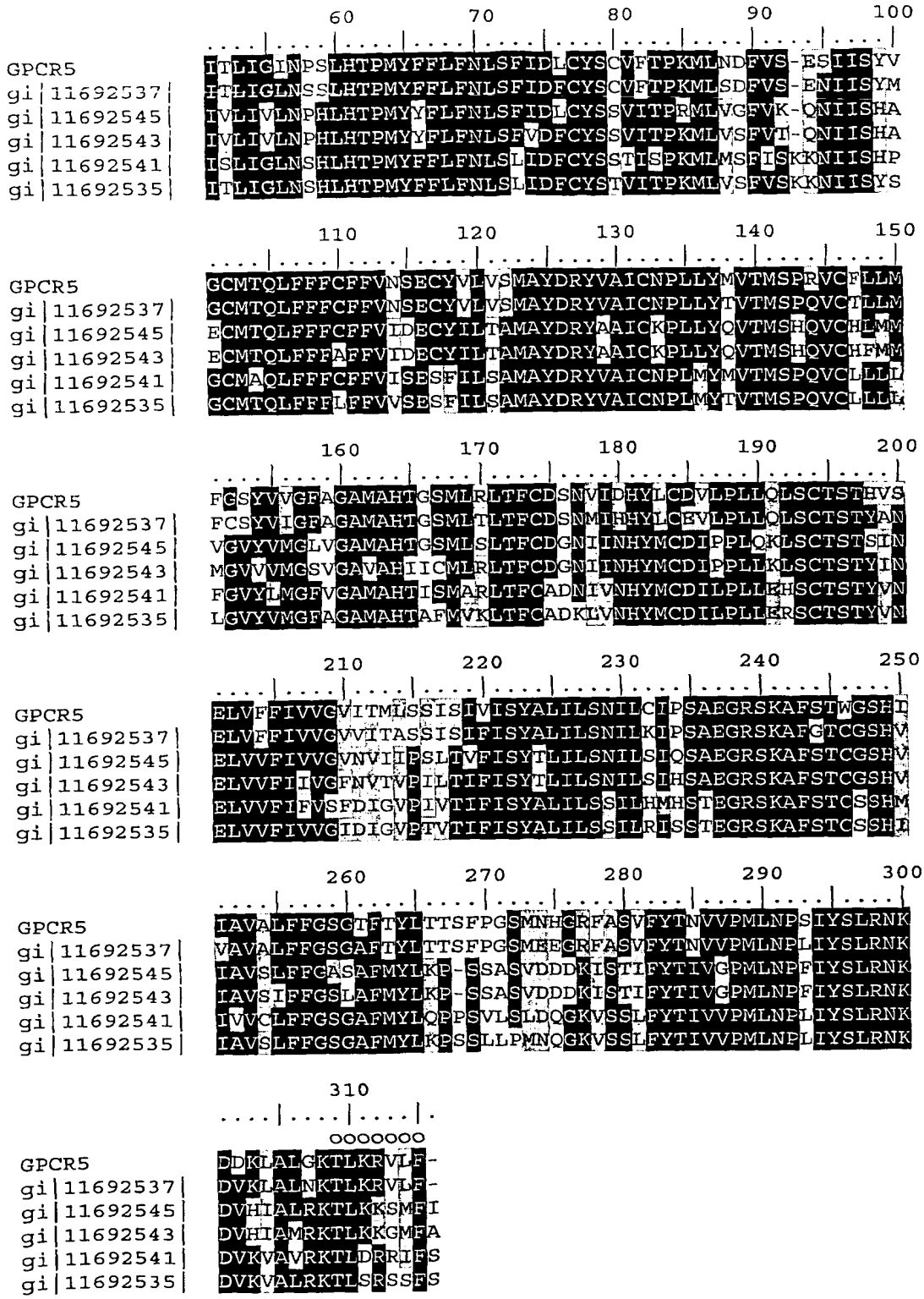
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|11692537|gb|AAG39865.1|AF282280_1 (AF282280)	odorant receptor K21h1 [Mus musculus]	314	256/309 (82%)	270/309 (86%)	e-113
gi|11692545|gb|AAG39869.1|AF282284_1 (AF282284)	odorant receptor K26 [Mus musculus]	314	196/309 (63%)	235/309 (75%)	7e-87
gi|11692543|gb|AAG39868.1|AF282283_1 (AF282283)	odorant receptor K25 [Mus musculus]	309	187/309 (60%)	229/309 (73%)	7e-85
gi|11692541|gb|AAG39867.1|AF282282_1 (AF282282)	odorant receptor K23 [Mus musculus]	310	195/310 (62%)	235/310 (74%)	2e-84
gi|11692535|gb|AAG39864.1|AF282279_1 (AF282279)	odorant receptor K21 [Mus musculus]	310	199/305 (65%)	235/305 (76%)	7e-84

A multiple sequence alignment is given in Table 5G, with the GPCR5 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR5 with related protein sequences, shown in Table 5F. The residue that differs between GPCR5a and GPCR5b is marked with the (o) symbol.

Table 5G. Information for the ClustalW proteins:

1. GPCR5b; SEQ ID NO:16
2. gi|11692537|gb|AAG39865.1|AF282280_1 odorant receptor K21h1 [Mus musculus]; SEQ ID NO:50
3. gi|11692545|gb|AAG39869.1|AF282284_1 odorant receptor K26 [Mus musculus]; SEQ ID NO:51
4. gi|11692543|gb|AAG39868.1|AF282283_1 odorant receptor K25 [Mus musculus]; SEQ ID NO:52
5. gi|11692541|gb|AAG39867.1|AF282282_1 odorant receptor K23 [Mus musculus]; SEQ ID NO:53
6. gi|11692535|gb|AAG39864.1|AF282279_1 odorant receptor K21 [Mus musculus]; SEQ ID NO:54

```
                                10        20        30        40        50
                        ....|....|....|....|....|....|....|....|....|....|
GPCR5                   -----MTLRNSSSVTEFILVGLSEQPELQLPLFLLFLGIYVFTVVGNLGL
gi|11692537|            MSQKRMAPRNSSSVTEFILVGFSNQPALQLPLFFVFLGIYVLTVIGNLGL
gi|11692545|            MMLGRMAFSNDSSVKEFILLGLTQQPELQMPLFFLFLGIYVVSMVGNLGL
gi|11692543|            -----MGFGNDSSVKEFILLGLTQQPELQLPLFFLFLGIYVVSIVGNLGL
gi|11692541|            -----MTAKN-SSVTEFILAGLTDQPGLRMPLFFLFLGFYMVTVVGNLGL
gi|11692535|            -----MATEN-ASVPEFILAGLTDQPGLRMPLFFLFLGFYMVTMVGNLGL
```

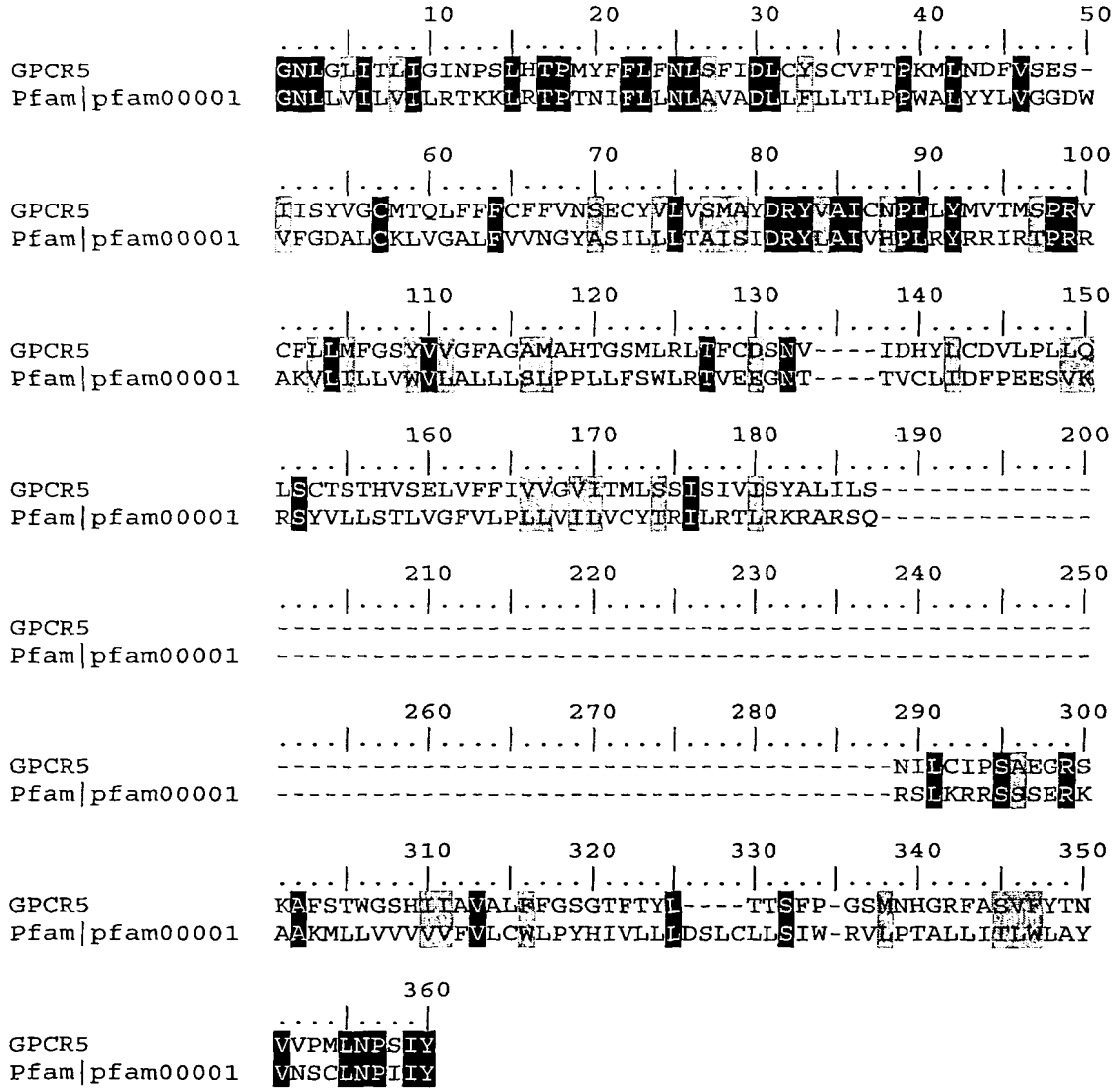


DOMAIN results for GPCR5 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 5H with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to

have significant homology to GPCR5. An alignment of GPCR5 residues 41-289 with 7tm_1 residues 1-254 (SEQ ID NO:34) are shown in Table 5H.

Table 5H. DOMAIN results for GPCR5

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	93.2	2e-20



The nucleic acids and proteins of GPCR5 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the

generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR5 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR5 epitope is from about amino acids 5 to 20. In other specific embodiments, GPCR5 epitopes are from about amino acids 230 to 245, from about amino acids 255 to 275 and from about amino acids 285 to 309.

GPCR6

A further GPCR-like protein of the invention, referred to herein as GPCR6, is an Olfactory Receptor ("OR")-like protein. The novel GPCR6 nucleic acid sequences were identified on chromosome 11 as described in Example 1. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR6 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR6 nucleic acids and encoded polypeptides are provided, namely GPCR6a and GPCR6b.

GPCR6a

In one embodiment, a GPCR6 variant is the novel GPCR6a (alternatively referred to herein as CG56025-01), which includes the 971 nucleotide sequence (SEQ ID NO:17) shown in Table 6A. A GPCR6a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 20-22 and ends with a TAA codon at nucleotides 956-958. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 6A, and the start and stop codons are in bold letters.

Table 6A. GPCR6 Nucleotide Sequence (SEQ ID NO:17)

TTCTAGGGTATCAAGGGAC**ATG**AGAAATGGCACAGTAATCACAGAATTCATCCTGCTAGGCTTTCCT
GTTATCCAAGGCCTACAAACACCTCTCTTTATTGCAATCTTTCTCACCTACATATTAACCCTTGCAG
GCAATGGGCTTATTATTGCCACTGTGTGGGCTGAGCCCAGGCTACAAATTCCAATGTACTTCTTCCT
TTGTAACTTGTCTTTCTTAGAAATCTGGTACACCACCACAGTCATCCCCAAACTGCTAGGAACCTTT
GTAGTGGCAAGAACAGTAATCTGCATGTCCTGCTGCCTGCTGCAGGCCTTCTTCCACTTCTTCGTGG
GCACCACCGAGTTCTTGATCCTCACTATCATGTCTTTTGACCGCTACCTCACCATCTGCAATCCCCT
TCACCACCCCACCATCATGACCAGCAAACTCTGCCTGCAGCTGGCCCTGAGCTCCTGGGTGGTGGGC
TTCACCATTGTCTTTTGTCAGACGATGCTGCTCATCCAGTTGCCATTCTGTGGCAATAATGTTATCA
GTCATTTCTACTGTGATGTTGGGCCCAGTTTGAAAGCCGCCTGCATAGACACCAGCATTTTGGAACT
CCTGGGCGTCATAGCAACCATCCTTGTGATCCCAGGGTCACTTCTCTTTAATATGATTTCTTATATC
TACATTCTGTCCGCAATCCTACGAATTCCTTCAGCCACTGGCCACCAAAAGACTTTCTCTACCTGTG
CCTCGCACCTGACAGTTGTCTCCCTGCTCTACGGGGCTGTTCTGTTCATGTACCTAAGACCCACAGC

ACACTCCTCCTTTAAGATTAATAAGGTGGTGTCTGTGCTAAATACTATCCTCACCCCCCTTCTGAAT
CCCTTTATTTATACTATTAGAAACAAGGAGGTGAAGGGAGCCTTAAGAAAGGCAATGACTTGCCCAA
AGACTGGTCATGCAAAG**TAA**AACATGCAACACA

The sequence of GPCR6a was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. The cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The cDNA coding for the GPCR6a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. The DNA sequence and protein sequence for a novel Olfactory Receptor-like gene were obtained by exon linking and are reported here as GPCR6a.These primers and methods used to amplify GPCR6 a cDNA are described in the Examples.

The GPCR6a polypeptide (SEQ ID NO:18) encoded by SEQ ID NO:17 is 312 aa in length, has a molecular weight of 34526.32 Daltons, and is presented using the one-letter amino acid code in Table 6B. The Psort profile for both GPCR6a and GPCR6b predicts that these sequences have a signal peptide and are likely to be localized at the endoplasmic reticulum (membrane) with a certainty of 0.685. In alternative embodiments, a GPCR6 polypeptide is located to the plasma membrane with a certainty of 0.6400, the Golgi body with a certainty of 0.4600, or to endoplamic reticulum (lumen) with a certainty of 0.1000. The Signal P predicts a likely cleavage site for a GPCR6 peptide is between positions 49 and 50, *i.e.*, at the dash in the sequence VWA-EP.

Table 6B. GPCR6a protein sequence (SEQ ID NO:18)

MRNGTVITEFILLGFPVIQGLQTPLFIAIFLTYILTLAGNGLIIATVWAEPRLQIPMYFFLCNLSFL
EIWYTTTVIPKLLGTFVVARTVICMSCCLLQAFFHFFVGTTEFLILTIMSFDRYLTICNPLHHPTIM
TSKLCLQLALSSWVVGFTIVFCQTMLLIQLPFCGNNVISHFYCDVGPSLKAACIDTSILELLGVIAT
ILVIPGSLLFNMISYIYILSAILRIPSATGHQKTFSTCASHLTVVSLLYGAVLFMYLRPTAHSSFKI
NKVVSVLNTILTPLLNPFIYTIRNKEVKGALRKAMTCPKTGHAK

GPCR6b

In an alternative embodiment, a GPCR6 variant is the novel GPCR6b (alternatively referred to herein as AP001804_B), which includes the 939 nucleotide sequence (SEQ ID NO:19) shown in Table 6C. The GPCR6b ORF begins with a Kozak consensus ATG

initiation codon at nucleotides 1-3 and ends with a TAA codon at nucleotides 937-939, which are in bold letters in Table 6C. The GPCR6 protein encoded by SEQ ID NO:7 is identical to SEQ ID NO:18, above.

Table 6C. GPCR6b Nucleotide Sequence (SEQ ID NO:19)

ATGAGAAATGGCACAGTAATCACAGAATTCATCCTGCTAGGCTTTCCTGTTATCCAAGGCC
TACAAACACCTCTCTTTATTGCAATCTTTCTCACCTACATATTAACCCTTGCAGGCAATGG
GCTTATTATTGCCACTGTGTGGGCTGAGCCCAGGCTACAAATTCCAATGTACTTCTTCCTT
TGTAACTTGTCTTTCTTAGAAATCTGGTACACCACCACAGTCATCCCCAAACTGCTAGGAA
CCTTTGTAGTGGCAAGAACAGTAATCTGCATGTCCTGCTGCCTGCTGCAGGCCTTCTTCCA
CTTCTTCGTGGGCACCACCGAGTTCTTGATCCTCACTATCATGTCTTTTGACCGCTACCTC
ACCATCTGCAATCCCCTTCACCACCCCACCATCATGACCAGCAAACTCTGCCTGCAGCTGG
CCCTGAGCTCCTGGGTGGTGGGCTTCACCATTGTCTTTTGTCAGACGATGCTGCTCATCCA
GTTGCCATTCTGTGGCAATAATGTTATCAGTCATTTCTACTGTGATGTTGGGCCCAGTTTG
AAAGCCGCCTGCATAGACACCAGCATTTTGGAACTCCTGGGCGTCATAGCAACCATCCTTG
TGATCCCAGGGTCACTTCTCTTTAATATGATTTCTTATATCTACATTCTGTCCGCAATCCT
ACGAATTCCTTCAGCCACTGGCCACCAAAAGACTTTCTCTACCTGTGCCTCGCACCTGACA
GTTGTCTCCCTGCTCTACGGGGCTGTTCTGTTCATGTACCTAAGACCCACAGCACACTCCT
CCTTTAAGATTAATAAGGTGGTGTCTGTGCTAAATACTATCCTCACCCCCCTTCTGAATCC
CTTTATTTATACTATTAGAAACAAGGAGGTGAAGGGAGCCTTAAGAAAGGCAATGACTTGC
CCAAAGACTGGTCATGCAAAG**TAA**

GPCR6 Clones

Unless specifically addressed as GPCR6a or GPCR6b, any reference to GPCR6 is assumed to encompass all variants. Residue differences between any GPCRX variant sequences herein are written to show the residue in the "a" variant, the residue position with respect to the "a" variant, and the residue in the "b" variant. For example, the GPCR6 nucleic acid sequences differ at the following position: G394T.

In a search of sequence databases, it was found, for example, that the GPCR6 nucleic acid sequence of this invention has 770 of 922 bases (83%) identical to a *Rattus norvegicus* Olfactory Receptor-like protein mRNA (GENBANK-ID:RNOLP4|acc:X80671). The full amino acid sequence of the protein of the invention was found to have 247 of 302 amino acid residues (81%) identical to, and 261 of 302 residues (86%) similar to, the 309 amino acid residue Olfactory Receptor-like protein from *Rattus norvegicus* (SPTREMBL-ACC:Q63395). In all BLAST herein, the "E-value" or "Expect" value is a numeric indication of the probability that the aligned sequences could have achieved their similarity to the BLAST query sequence by chance alone, within the database that was searched. In addition, it was found, for example, that the GPCR6 nucleic acid sequence has 778 of 927 bases (83%) identical to a gb:GENBANK-ID:AF282289| acc:AF282289.1 mRNA from Mus musculus (Mus musculus odorant receptor K40 gene, complete cds). The full amino acid sequence of the

protein of the invention was found to have 259 of 304 amino acid residues (85%) identical to, and 275 of 304 amino acid residues (90%) similar to, the 308 amino acid residue ptnr:TREMBLNEW-ACC:AAG39874 protein from Mus musculus (Mouse) (ODORANT RECEPTOR K40). Additional BLAST results are shown in Table 6E.

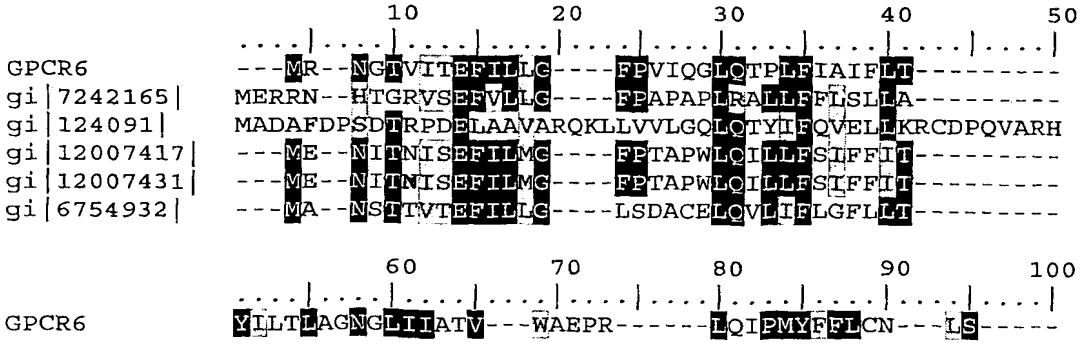
Table 6E. BLAST results for GPCR6

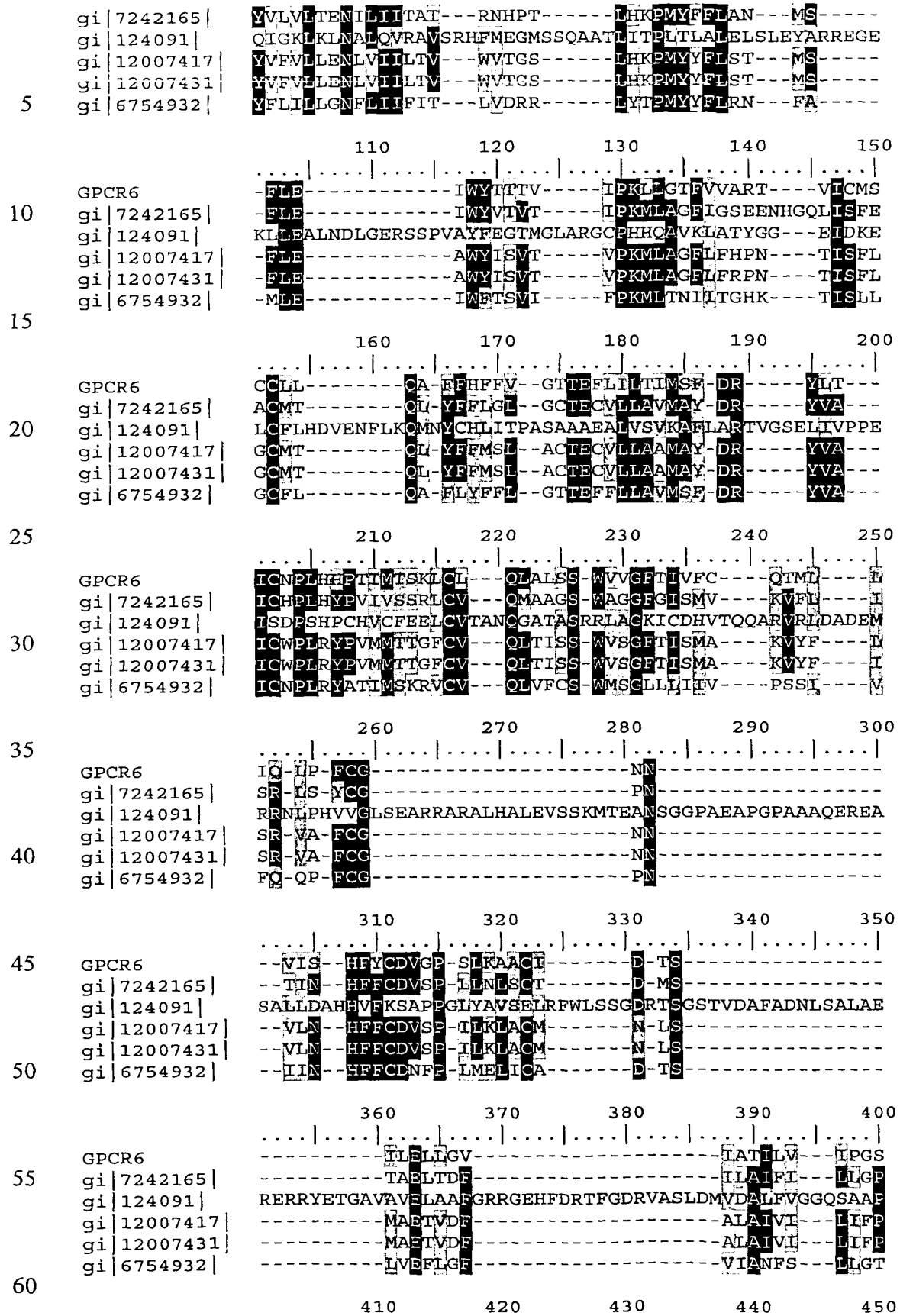
Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|7242165|ref|NP_035113.1|; gi|3983437|gb|AAD13307.1| (AF106007); gi|12007413|gb|AAG45187.1| (AF321233)	olfactory receptor 41 [Mus musculus]; olfactory receptor I7 [Mus musculus]	327	143/301 (47%)	193/301 (63%)	4e-67
gi|129091|sp|P23267|; gi|112091|pir||C23701; gi|205818|gb|AAA41741.1| (M64378)	OLF6_RAT OLFACTORY RECEPTOR-LIKE PROTEIN F6 [Rattus norvegicus]	311	142/294 (48%)	202/294 (68%)	5e-67
gi|12007417|gb|AAG45190.1| (AF321234)	m50 olfactory receptor [Mus musculus]	316	144/303 (47%)	200/303 (65%)	1e-65
gi|12007431|gb|AAG45202.1|AF321236_1 (AF321236)	m50 olfactory receptor [Mus musculus]	316	143/303 (47%)	200/303 (65%)	2e-65
gi|6754932|ref|NP_035121.1|; gi|3983374|gb|AAD13315.1| (AF102523)	olfactory receptor 49 [Mus musculus]; olfactory receptor C6 [Mus musculus]	313	142/302 (47%)	196/302 (64%)	4e-65

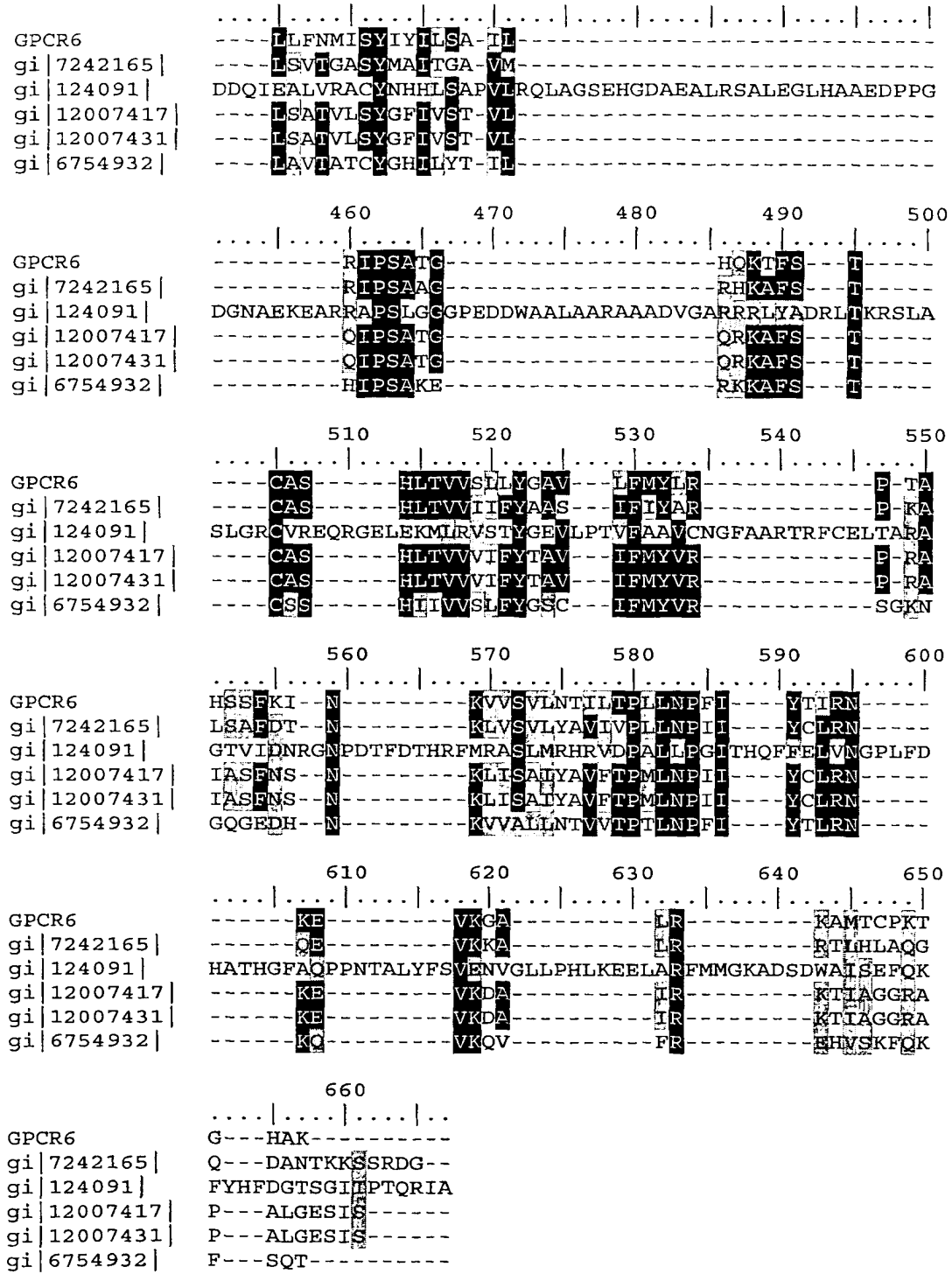
A multiple sequence alignment is given in Table 6F, with the GPCR6 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR6 with related protein sequences.

Table 6F. Information for the ClustalW proteins:

1. GPCR6; SEQ ID NO:18
2. gi|7242165|ref|NP_035113.1| olfactory receptor 41 [Mus musculus]; SEQ ID NO:55
3. gi|129091|sp|P23267|OLF6_ rat olfactory receptor - Like Protein F6; SEQ ID NO:56
4. gi|12007417|gb|AAG45190.1| (AF321234) m50 olfactory receptor [Mus musculus]; SEQ ID NO:57
5. gi|12007431|gb|AAG45202.1|AF321236_1 m50 olfactory receptor [Mus musculus]; SEQ ID NO:58
6. gi|6754932|ref|NP_035121.1| olfactory receptor 49 [Mus musculus]; SEQ ID NO:59




gi|7242165|
gi|124091|
gi|12007417|
gi|12007431|
gi|6754932|
GPCR6
110 120 130 140 150
160 170 180 190 200
210 220 230 240 250
260 270 280 290 300
310 320 330 340 350
360 370 380 390 400
410 420 430 440 450

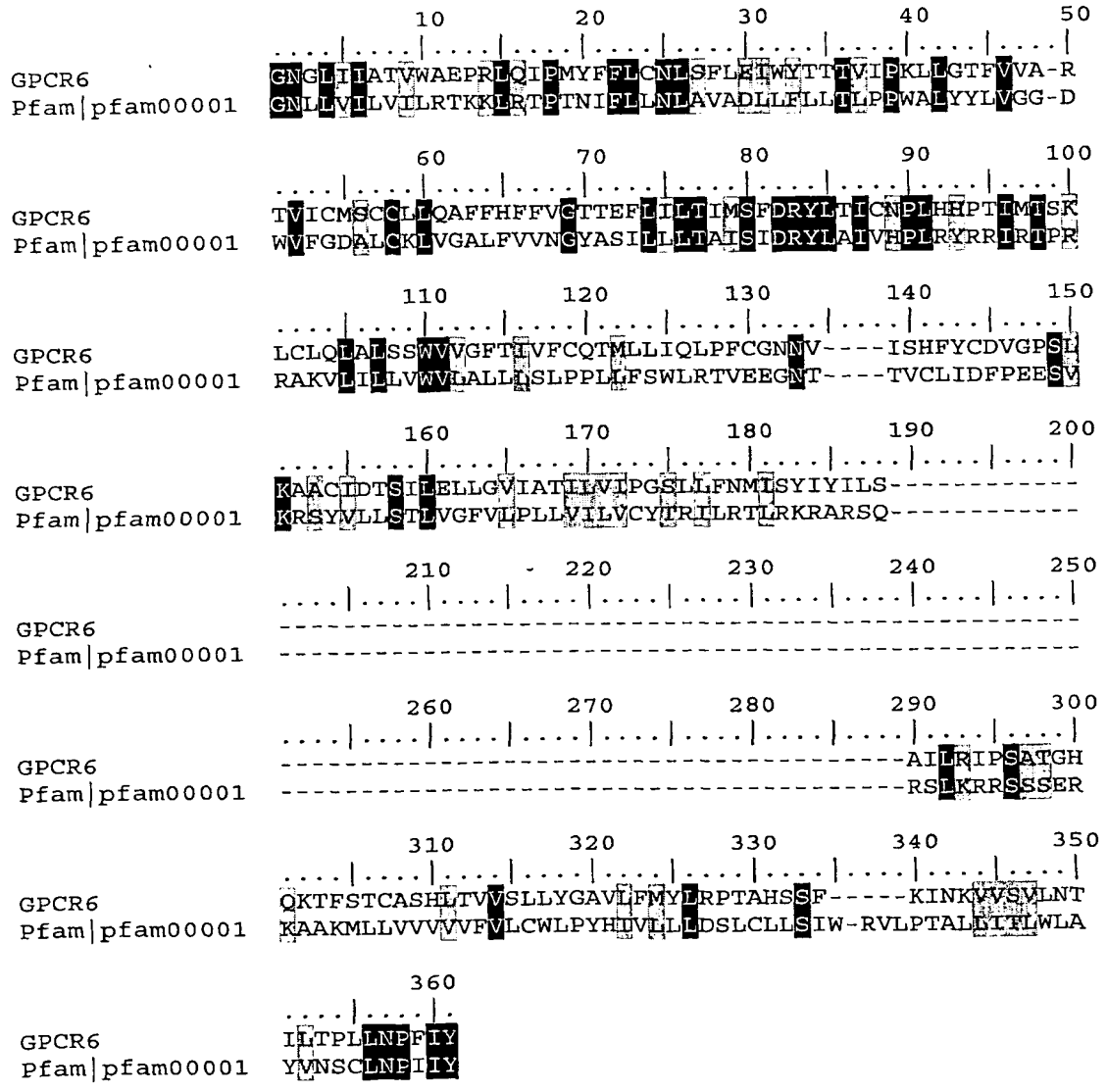


DOMAIN results for GPCR6 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 6G with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to

have significant homology to GPCR6. An alignment of GPCR6 residues 39-288 (SEQ ID NO:18) with 7tm_1 residues 1-254 (SEQ ID NO:34) are shown in Table 6G.

Table 6G. DOMAIN results for GPCR6

PSSMs producing significant alignments:		Score (bits)	E value
gnl|Pfam|pfam00001	7tm_1, 7 transmembrane receptor (rhodopsin family)	88.2	6e-19



The Olfactory Receptor-like GPCR6 proteins disclosed are expressed in at least the following tissues: olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types especially olfactory epithelium. Further tissue expression analysis is provided in the Examples.

The nucleic acids and proteins of GPCR6 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR6 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR6 epitope is from about amino acids 115 to 135. In additional embodiments, GPCR6 epitopes are from about amino acids 225 to 240 and from about amino acids 280 to 305.

GPCR7

The disclosed novel GPCR7 nucleic acid of 942 nucleotides (also referred to as 20722608_EXT) is shown in Table 7A. An ORF begins with an ATG initiation codon at nucleotides 1-3 and ends with a TGA codon at nucleotides 940-942.

The following genomic clone was identified as having regions of high homology to Olfactory Receptors of the invention: genomic clone >acc:AP000868 HTG *Homo sapiens* chromosome 11 clone RP11-688B18 map 11q24, WORKING DRAFT SEQUENCE, in unordered pieces - *Homo sapiens*, 176200 bp (DNA). The sequence was analyzed by GENSCAN and GRAIL software programs to identify exons and putative coding sequences. The start and stop codons in Table 7A are in bold letters.

Table 7A. GPCR7 Nucleotide Sequence (SEQ ID NO:20)

ATGGGAAACTGGAGCACTGTGACTGAAATCACCCTAATTGCCTTCCCAGCTCTCCTGGAGA
TTCGAATATCTCTCTTCGTGGTTCTTGTGGTAACTTACACATTAACAGCAACAGGAAACAT
CACCATCATCTCCCTGATATGGATTGATCATCGCCTGCAAACTCCAATGTACTTCTTCCTC
AGTAATTTGTCCTTTCTGGATATCTTATACACCACTGTCATTACCCCAAAGTTGTTGGCCT
GCCTCCTAGGAGAAGAGAAAACCATATCTTTTGCTGGTTGCATGATCCAAACATATTTCTA
CTTCTTTCTGGGGACGGTGGAGTTTATCCTCTTGGCGGTGATGTCCTTTGACCGCTACATG
GCTATCTGCGACCCACTGCACTACACGGTCATCATGAACAGCAGGGCCTGCCTTCTGCTGG
TTCTGGGATGCTGGGTGGGAGCCTTCCTGTCTGTGTTGTTTCCAACCATTGTAGTGACAAG
GCTACCTTACTGTAGGAAAGAAATTAATCATTTCTTCTGTGACATTGCCCCTCTTCTTCAG
GTGGCCTGTATAAATACTCACCTCATTGAGAAGATAAACTTTCTCCTCTCTGCCCTTGTCA
TCCTGAGCTCCCTGGCATTCACTACTGGGTCCTACGTGTACATAATTTCTACCATCCTGCG
TATCCCCTCCACCCAGGGCCGTCAGAAAGCTTTTTCTACCTGTGCTTCTCACATCACTGTT
GTCTCCATTGCCCACGGGAGCAACATCTTTGTGTATGTGAGACCCAATCAGAACTCCTCAC
TGGATTATGACAAGGTGGCCGCTGTCCTCATCACAGTGGTGACCCCTCTCCTGAACCCTTT
TATCTACAGCTTGAGGAATGAGAAGGTACAGGAAGTGTTGAGAGAGACAGTGAACAGAATC

ATGACCTTGATACAAAGGAAAACTTGA

The GPCR7 protein (SEQ ID NO:21) encoded by SEQ ID NO:20 has 313 amino acid residues and is presented using the one-letter code in Table 7B. The predicted molecular weight of GPCR7 protein is 35326.06 Daltons. The Psort profile for GPCR7 predicts that this sequence has a signal peptide and is likely to be localized at the endoplasmic reticulum (membrane) with a certainty of 0.685. In alternative embodiments, GPCR7 is located in the plasma membrane with a certainty of 0.676, the Golgi body with a certainty of 0.460 or endoplasmic reticulum (lumen) with a certainty of 0.100. The Signal P predicts a likely cleavage site between positions 36 and 37, i.e., at the dash in the sequence TLT-AT.

Table 7B. Encoded GPCR7 protein sequence (SEQ ID NO:21)

MGNWSTVTEITLIAFPALLEIRISLFVVLVVTYTLTATGNITIISLIWIDHRLQTPMYFF
LSNLSFLDILYTTVITPKLLACLLGEEKTISFAGCMIQTYFYFFLGTVEFILLAVMSFDR
YMAICDPLHYTVIMNSRACLLLVLGCWVGAFLSVLFPTIVVTRLPYCRKEINHFFCDIAP
LLQVACINTHLIEKINFLLSALVILSSLAFTTGSYVYIISTILRIPSTQGRQKAFSTCAS
HITVVSIAHGSNIFVYVRPNQNSSLDYDKVAAVLITVVTPLLNPFIYSLRNEKVQEVLRE
TVNRIMTLIQRKT

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of this invention 20722608_EXT has 581 of 916 bases (63%) identical to a *Mus musculus* Olfactory Receptor-like protein mRNA (GENBANK-ID:AF102523). The full amino acid sequence of the protein of the invention was found to have 149 of 312 amino acid residues (47%) identical to, and 211 of 312 residues (67%) similar to, the 313 amino acid residue Olfactory Receptor-like protein from *Mus musculus* (SPTREMBL-ACC:Q9Z1V0). GPCR7 also has homology to the proteins shown in the BLASTP data in Table 7C.

Table 7C. BLAST results for GPCR7

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|129091|sp|P23267|; gi|112091|pir||C23701; gi|205818|gb|AAA41741.1| (M64378)	OLF6_RAT OLFACTORY RECEPTOR-LIKE PROTEIN F6 [Rattus norvegicus]	311	141/301 (46%)	184/301 (60%)	1e-63
gi|6754932|ref|NP_035121.1|; gi|3983374|gb|AAD13315.1| (AF102523)	olfactory receptor 49 [Mus musculus]; olfactory receptor C6 [Mus musculus]	313	142/312 (45%)	195/312 (61%)	2e-63

gi|7242165|ref|NP_035113.1|; gi|3983437|gb|AAD13307.1| (AF106007); gi|12007413|gb|AAG45187.1| (AF321233)	olfactory receptor 41; olfactory receptor I7 [Mus musculus]	327	133/302 (44%)	181/302 (59%)	5e-62
gi|10181106|ref|NP_065623.1|; gi|7638409|gb|AAF65461.1|AF247657_1 (AF247657); gi|12007410|gb|AAG45184.1| (AF321233)	olfactory receptor 17; olfactory receptor P2; P2 olfactory receptor [Mus musculus]	315	133/307 (43%)	184/307 (59%)	2e-61
gi|13928994|ref|NP_113898.1|; gi|129092|sp|P23270|; gi|112099|pir||F2370|; gi|205834|gb|AAA41749.1| (M64386)	olfactory receptor 41; OLF7_RAT olfactory receptor -like protein I7 olfactory receptor I7 - rat [Rattus norvegicus]	327	131/302 (43%)	179/302 (58%)	2e-60

A multiple sequence alignment is given in Table 7D, with the protein of the invention 20722608_EXT being shown on line 1, in a ClustalW analysis comparing the protein of the invention with related protein sequences. This BLASTP data is displayed graphically in the ClustalW in Table 7D.

Table 7D. ClustalW Analysis of GPCR7

1) GPCR7; SEQ ID NO:21
2)>gi|129091|sp|P23267|OLF6_rat olfactory receptor - Like Protein F6; SEQ ID NO:60
3) >gi|6754932|ref|NP_035121.1| olfactory receptor 49 [Mus musculus] ; SEQ ID NO:61
4) >gi|7242165|ref|NP_035113.1| olfactory receptor 41 [Mus musculus] ; SEQ ID NO:62
5) >gi|10181106|ref|NP_065623.1| olfactory receptor 17 [Mus musculus] ; SEQ ID NO:63
6) >gi|13928994|ref|NP_113898.1| olfactory receptor 41 [Rattus norvegicus] ; SEQ ID NO:64

```
                                10        20        30        40        50
                    ....|....|....|....|....|....|....|....|....|....|
GPCR7               -----MGNWSTVTEITLIAFP-ALLEIRISLFVVLVVTYTLTATGNITII
gi|129091|          MAWSTGQNLSTPGPFILLGFP-GPRSMRIGLFLLFLVMYLLTVVGNLAII
gi|6754932|         -----MANSTTVTEFILLGLS-DACELQVLIFLGFLLTYFLILLGNFLII
gi|7242165|         --MERRNHTGRVSEFVLLGFP-APAPLRALLFFLSLLAYVLVLTENILII
gi|10181106|        ---MTWGNWTTVREFILMSFSSLSYEVQALLFLLFLIIYLVTLMGNVLII
gi|13928994|        --MERRNHSGRVSEFVLLGFP-APAPLRVLLFFLSLLXYVLVLTENMLII

                                60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|
GPCR7               SLIWIDHRLQTPMYFFLSNLSFLDILYTTVITPKLLACLLG----EEKTI
gi|129091|          SLVGAHRCLQTPMYFFLCNLSFLEIWFTTACVPKTLATFAP----RGGVI
gi|6754932|         FITLVDRRLYTPMYYFLRNFAMLEIWFTSVIFPKMLTNIIT----GHKTI
gi|7242165|         TAIRNHPTLHKPMYFFLANMSFLEIWYVTVTIPKMLAGFIGSEENHGQLI
gi|10181106|        LVTTADSALQSPMYFFLRNLSFLEIGFNLVIVPKMLSTLIL----QDKTI
gi|13928994|        IAIRNHPTLHKPMYFFLANMSFLEIWYVTVTIPKMLAGFIGSKENHGQLI

                               110       120       130       140       150
                    ....|....|....|....|....|....|....|....|....|....|
```

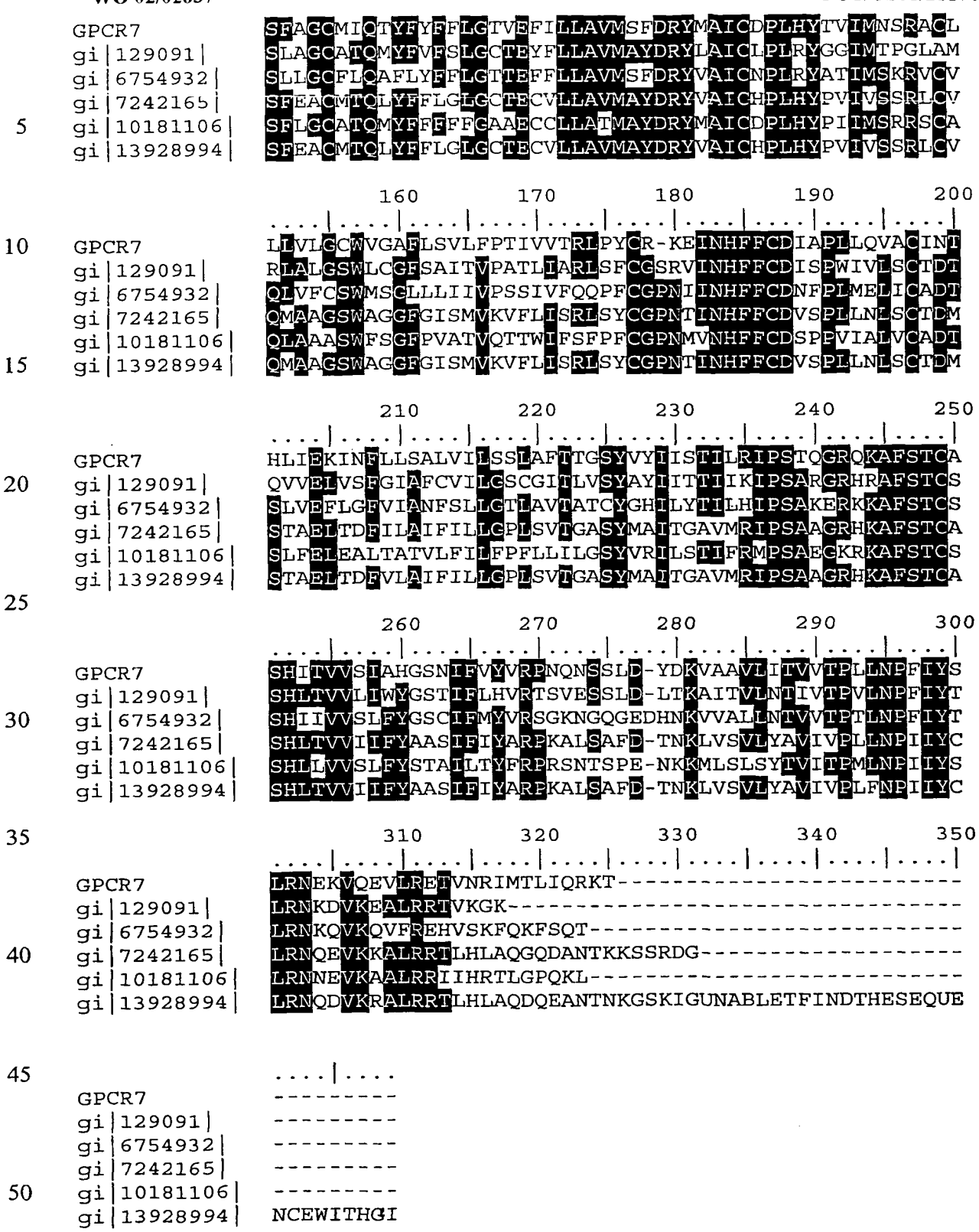


Table 7E lists the domain description from DOMAIN analysis results against GPCR7. This indicates that the GPCR7 sequence has properties similar to those of other proteins

known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34) itself.

Table 7E Domain Analysis of GPCR7

PSSMs producing significant alignments: Score (bits) E value

gnl|Pfam|pfam00001 7tm_1, 7 transmembrane receptor (rhodopsin family) 108 4e-25

```
                    10        20        30        40        50
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               GNITIISLIWIDHRLQ--TPMYFFLSNLSFLDILYTTVIIPKLLACLLGE
Pfam|pfam00001      GNLLVILVILRTKKLR--TPTNIFLLNLAVADLLFLLTLPFWALYYLVGG

                    60        70        80        90        100
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               -EKTISFAGCMIQTYFYFFLGTVEFILLAVMSFDRYMAICDPLHYTVIMN
Pfam|pfam00001      -DWVFGDALCKLVGALFVVNGYASILLLTAISIDRYLAIVHPLRYRRIRT

                    110       120       130       140       150
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               SRACLLLVLGCWVGAFLSVLFPTIVVTRLPYCRKEINH-----FFCDIAP
Pfam|pfam00001      PRRAKVLILLVWVLALLLSLPPLLFSWLRTV--EEGNT-----TVCLIDF

                    160       170       180       190       200
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               LLQVAC-------INTHLIEKINFLLSALVILSSLAFTTG---SYVYIIS
Pfam|pfam00001      PEESVK-------RSYVLLSTLVGFVLPLLVILVCYTRILRTLRKRARSQ

                    210       220       230       240       250
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               --------------------------------------------------
Pfam|pfam00001      --------------------------------------------------

                    260       270       280       290       300
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               --------------------------------------------------
Pfam|pfam00001      --------------------------------------------------

                    310       320       330       340       350
                    |....|....|....|....|....|....|....|....|....|....|
GPCR7               -TILRIPSTQGRQKAFSTCASHITVVSIAHGSNIFVYVRPNQNSS-----
Pfam|pfam00001      -RSLKRRSSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIW-RV

                    360       370
                    |....|....|....|...
GPCR7               LDYDKVAAVLITVVTPLLNPFIY
Pfam|pfam00001      LPTALLITLWLAYVNSCLNPIIY
```

The GPCR7 protein disclosed in this invention is expressed in at least the following human tissues: pancreas and olfactory epithelium; This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types.

The nucleic acids and proteins of GPCR7 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein.

The novel GPCR7 nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR7 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR7 epitope is from about amino acids 120 to 132. In additional embodiments, GPCR7 epitopes are from about amino acids 160 to 180, from about amino acids 215 to 240, from about amino acids 255 to 270 and from about amino acids 280 to 310.

GPCR8

The disclosed novel GPCR8 nucleic acid of 1920 nucleotides (also referred to as CG-SC931712) is shown in Table 8A. An ORF begins with an ATG initiation codon at nucleotides 201-203 and ends with a TGA codon at nucleotides 1137-1139.

The following genomic clone was identified as having high homology to olfactory receptor-like protein (HS6M1-6). The start and stop codons in Table 8A are in bold letters and the putative untranslated regions upstream from the initiation site and downstream from the termination codon are underlined.

Table 8A. GPCR8 Nucleotide Sequence (SEQ ID NO:22)

ATTGCTTGTATCTGTAGCTGGATAAATATCTCAATGAAGCATATAAAGGGAACTGTATAAAAATTC
TACTACCATTATGGTGCACACTCTCTGGAAGTGGGATACTTTTGTCTTCAATCTGTTTGCAAGTGA
GCGGTTGACAATGCATGGACAGACTTTGAGTTTATGTGGTTCTTTCTTTAGGTATAAGAAAAAGAT
GAATGATGATTAAAAAAAATGCAAGTTCGGAAGACTTCTTTATTCTACTTGGATTTTCTAATTGGC
CTCAGCTGGAAGTAGTTCTCTTTGTGGTTATCTTGATCTTCTACCTGATGACACTGACAGGAAACC
TGTTCATCATCATCCTGTCATACGTGGACTCCCATCTCCACACACCAATGTACTTCTTCCTTTCAA
ACCTCTCATTTCTGGATCTCTGCCACACCACCAGCTCTATCCCTCAGTTGCTGGTGAATCTCCGGG
GCCCGGAAAAGACCATCTCGTATGCTGGTTGCATGGTTCAACTTTACTTTGTTCTTGCACTGGGAA
TCGCAGAGTGTGTCCTACTGGTGGTGATGTCGTATGATCGTTATGTAGCTGTGTGTAGACCTTTGC
ATTACACTGTCCTCATGCACCCTCGTTTCTGCCACTTGTTGGCTGCGGCTTCTTGGGTAATTGGTT
TTACTATCTCAGCACTTCATTCCTCCTTTACTTTCTGGGTACCCCTTTGTGGACATCGCCTAGTGG
ATCACTTCTTCTGTGAAGTTCCAGCACTTCTGCGTTTATCATGTGTTGACACCCATGCAAATGAGC

TGACCCTCATGGTCATGAGCTCCATTTTTGTTCTCATACCTCTCATTCTGATTCTCACTGCCTATG
GTGCCATTGCCCGGGCTGTACTGAGCATGCAATCAACCACTGGGCTTCAGAAAGTGTTTAGGACAT
GTGGAGCCCATCTTATGGTTGTATCTCTCTTTTTCATTCCAGTCATGTGCATGTATCTCCAGCCAC
CATCAGAAAATTCTCCTGATCAGGGCAAGTTCATTGCCCTCTTTTATACTGTTGTCACACCGAGTC
TTAATCCTCTAATCTACACTCTCAGAAACAAGCATGTAAAAGGGGCAGCGAAGAGACTATTGGGGT
GGGAGTGGGGGAAG**TGA**CAGGGAAATCATGTTGTCTGTTGTCATTGTTTTTCCTAGGGTCTTAGCC
ATCTTGAAAGGTGGTTTCCCTGCTTCTTTGTGATTTATTTTTGTTCTAACAGCTCACAAAACATGG
AATAGTTCAGTTCCCCCATTTGTTGCTCTGTTTAATATTTAGTTCTGAAATATTATGTTGAGATAA
AGGTTTTGATTAGTACCATTTTGTTCTTTTACAATTGTATATTTTATTTCTGTGAAAATTGTGGAC
TGTGGTTTCAACGTAAATAAATGTGCATGCGAATAGTTATGAGGAGATTATTTAAAAAATATTGGC
AATATTTCTGACAATGTGCTAAATTATGAACTGACCATTGATATGTATAGGAAGAGAAGGGCAATA
TTGCAAAGATGTAGGCTGAAGAAGTTTTTGGTTATTAAATAAACCTTAAATGAAGCTAAAAATAGT
CACAGCAAAGAAAAATAGTAAACATAATGAATAACACCATTTATTATATGGTAAAGGATATGTCAT
AATTTTTTGGTTGAAGTTCACTTTTTAAAGACACTAAATTATATAATTTATCCTGTAGGTCTGCAT
TCTTGTCACATTGAACAGTAAACTAATATCTCTTTAAAATGGCTGATTCGTTCATCTGTCCATTTA
TTCATTAACTTATTCTTCATTAGCTAAATCTTACTGGACATGTACTCTCTCCCAGTTTGTGAAATT
CTTGGTAACATGTATAAATATAACATACTTTGTCTGAACAGAATGCACTCTCTATCGGGAAAAATG
GCAACA

The GPCR8 protein (SEQ ID NO:23) encoded by SEQ ID NO:22 has 312 amino acid residues and is presented using the one-letter code in Table 8B. The predicted molecular weight of GPCR8 protein is 35202.21 Daltons. The GPCR8 amino acid sequence is 100% homologous to olfactory receptor-like protein (HS6M1-6). The Psort profile for GPCR8 predicts that this sequence has a signal peptide and is likely to be localized at the endoplasmic reticulum (membrane). In an alternative embodiment, GPCR8 is located in the plasma membrane.

Table 8B. Encoded GPCR8 protein sequence (SEQ ID NO:23)

MMIKKNASSEDFFILLGFSNWPQLEVVLFVVILIFYLMTLTGNLFIIILSYVDSHLHTPMYFFLSNLS
FLDLCHTTSSIPQLLVNLRGPEKTISYAGCMVQLYFVLALGIAECVLLVVMSYDRYVAVCRPLHYTVL
MHPRFCHLLAAASWVIGFTISALHSSFTFWVPLCGHRLVDHFFCEVPALLRLSCVDTHANELTLMVMS
SIFVLIPLILILTAYGAIARAVLSMQSTTGLQKVFRTCGAHLMVVSLFFIPVMCMYLQPPSENSPDQG
KFIALFYTVVTPSLNPLIYTLRNKHVKGAAKRLLGWEWGK

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of this invention CG-SC931712 is greater than 95% homologous to olfactory receptor-like protein (HS6M1-6) and gi|13624331 GPCR8 also has homology to the proteins shown in the BLASTP data in Table 8C.

Table 8C. BLAST results for GPCR8

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect

gi|13624331|ref|NP_112167.1| ; gi|14423776|sp|O76002| ; gi|3757728|emb|CAA18784.1| (AL022727); gi|12054367|emb|CAC20491.1| (AJ302571)	olfactory receptor, family 2, subfamily J, member 2 [Homo sapiens]; O2J2_HUMAN OLFACTORY RECEPTOR 2J2 (OLFACTORY RECEPTOR 6-8) (OR6-8) (HS6M1-6); dJ80I19.4 (olfactory receptor-like protein (hs6M1-6)); olfactory receptor [Homo sapiens]	312	302/312 (96%)	302/312 (96%)	e-154
gi|12054379|emb|CAC20497.1| (AJ302577); through gi|12054389|emb|CAC20502.1| (AJ302582)	olfactory receptor [Homo sapiens]	312	299/312 (95%)	300/312 (95%)	e-152
gi|12054391|emb|CAC20503.1| (AJ302583)	olfactory receptor [Homo sapiens]	312	298/312 (95%)	299/312 (95%)	e-151
gi|12054355|emb|CAC20485.1| (AJ302565); through gi|12054365|emb|CAC20490.1| (AJ302570)	olfactory receptor [Homo sapiens]	312	275/311 (88%)	284/311 (90%)	e-139
gi|12054359|emb|CAC20487.1| (AJ302567); gi|12054361|emb|CAC20488.1| (AJ302568)	olfactory receptor [Homo sapiens]	312	274/311 (88%)	284/311 (91%)	e-139

A multiple sequence alignment is given in Table 8D, with the protein of the invention 20722608_EXT being shown on line 1, in a ClustalW analysis comparing the protein of the invention with related protein sequences, shown in Table 8C. This BLASTP data is displayed graphically in the Clustal W in Table 8D.

Table 8D. ClustalW Analysis of GPCR8

1) GPCR8; SEQ ID NO:23
2) >gi|13624331|ref|NP_112167.1| olfactory receptor, family 2, subfamily J, member 2 [Homo sapiens]; SEQ ID NO:65
3) >gi|12054379|emb|CAC20497.1| AJ302577 olfactory receptor [Homo sapiens]; SEQ ID NO:66
4) >gi|12054391|emb|CAC20503.1| AJ302583 olfactory receptor [Homo sapiens]; SEQ ID NO:67
5) >gi|12054355|emb|CAC20485.1| AJ302565 olfactory receptor [Homo sapiens]; SEQ ID NO:68
6) >gi|12054359|emb|CAC20487.1| AJ302567 olfactory receptor [Homo sapiens]; SEQ ID NO:69

```
                    10        20        30        40        50
                    |....|....|....|....|....|....|....|....|....|....|
GPCR8               MMIKKNASSEDFFILLGFSNWPQLEVVLFVVILIFYLMTLTGNLFIIILS
gi|13624331|        MMIKKNASSEDFFILLGFSNWPQLEVVLFVVILIFYLMTLTGNLFIIILS
gi|12054379|        MMIKKNASSEDFFILLGFSNWPQLEVVLFVVILIFYLMTLTGNLFIIILS
gi|12054391|        MMIKKNASSEDFFILLGFSNWPQLEVVLFVVILIFYLMTLTGNLFIIILS
gi|12054355|        MLMKKNASFEDFFILLGFSNWPHLEVVLFVVILIFYLITLIGNLFIIILS
gi|12054359|        MLMKKNASFEDFFLLLGFSNWPHLEVVLFVVILIFYLITLIGNLFIIILS

                    60        70        80        90        100
                    |....|....|....|....|....|....|....|....|....|....|
GPCR8               YVDSHLHTPMYFFLSNLSFLDLCHTTSSIPQLLVNLRGPEKTISYAGCMV
gi|13624331|        YVDSHLHTPMYFFLSNLSFLDLCHTTSSIPQLLVNLRGPEKTISYAGCMV
gi|12054379|        YVDSHLHTPMYFFLSNLSFLDLCYTTSSIPQLLVNLRGPEKTISYAGCMV
```

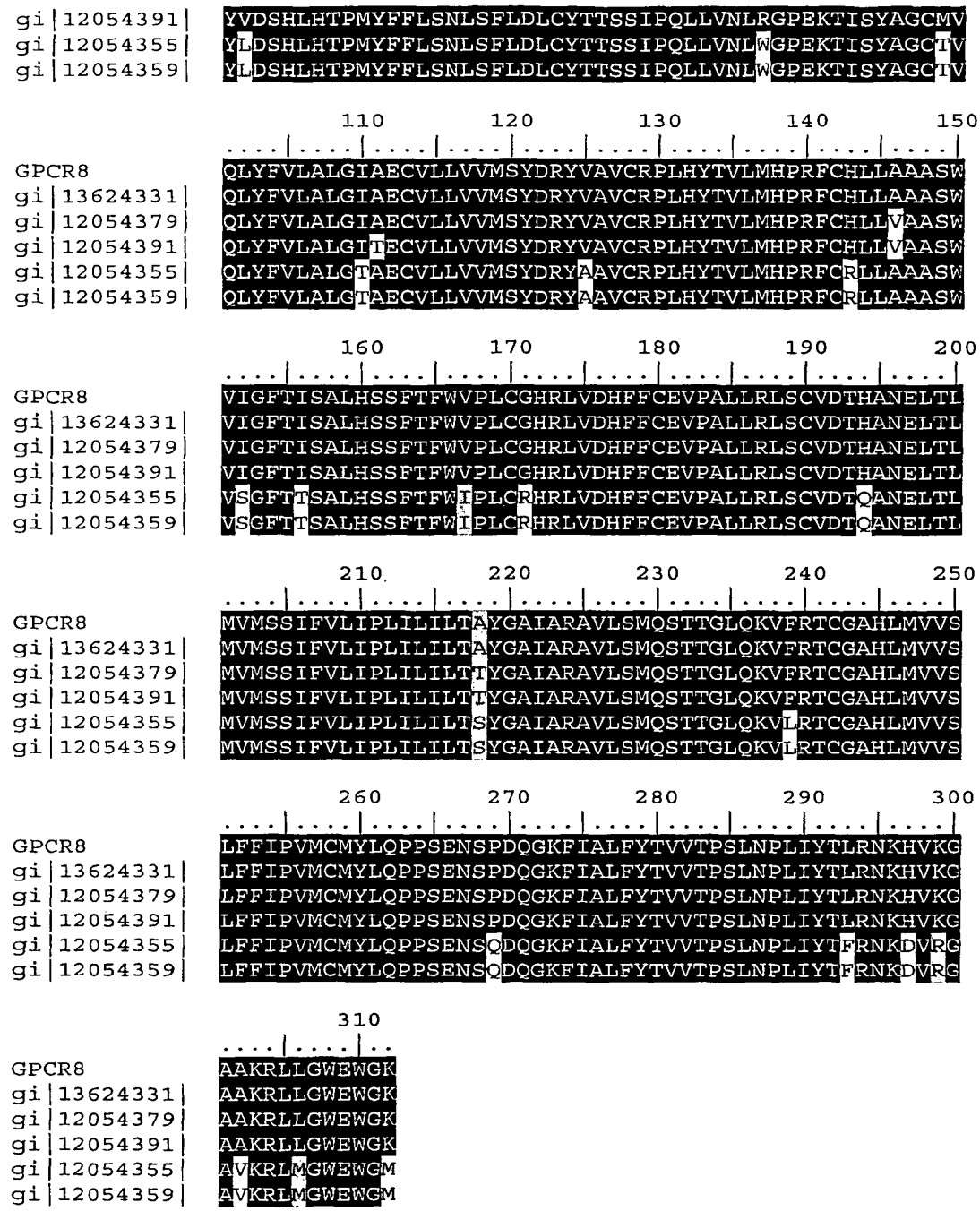
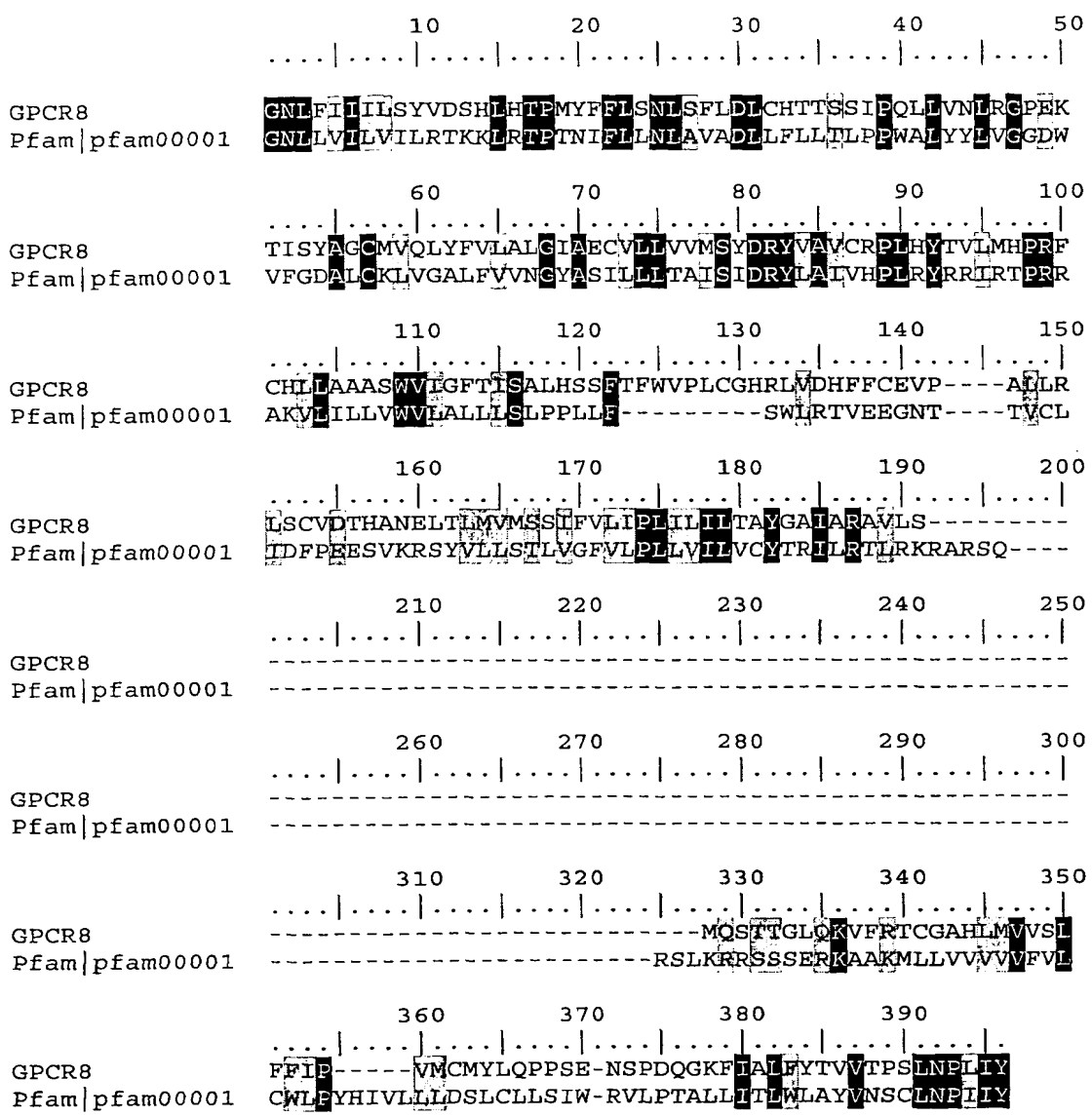


Table 8E lists the domain description from DOMAIN analysis results against GPCR8. This indicates that the GPCR8 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34) itself.

Table 8E Domain Analysis of GPCR8

PSSMs producing significant alignments:	Score	E



The Olfactory Receptor-like protein disclosed in this invention is expressed in at least the following human tissues: pancreas and olfactory epithelium; This is by no way limiting in that olfactory receptors are a class of G protein-coupled receptor which are known to be expressed in all tissue types.

The GPCR8 nucleic acids and proteins are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein. A monoclonal antibody targeting CG-SC931712 protein, specifically its extracellular region, will have a therapeutic role in treating cancer. It will also

have a role in treating angiogenesis related diseases. Being a GPCR, it could be used to screen for small molecule drug to treat cancer.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR8 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR8 epitope is from about amino acids 1 to 15. In additional embodiments, GPCR8 epitopes are from about amino acids 80 to 95, from about amino acids 115 to 130, from about amino acids 165 to 175, from about amino acids 180 to 195, from about amino acids 230 to 245, from about amino acids 255 to 270 and from about amino acids 285 to 305.

GPCR9

A second GPCR-like protein of the invention, referred to herein as GPCR9, is an Olfactory Receptor ("OR")-like protein. The GPCR9 gene maps to chromosome 9 p13.1-13.3. Some members of the Olfactory Receptor-Like Protein Family end up localized at the cell surface, where they exhibit activity. Therefore it is likely that these novel GPCR9 proteins are available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Two alternative novel GPCR9 nucleic acids, namely GPCR9a and GPCR9b, and an encoded GPCR9 polypeptide are provided,.

GPCR9a

In one embodiment, a GPCR9 variant is the novel GPCR9a (alternatively referred to herein as 21629632.0.20), which includes the 2028 nucleotide sequence (SEQ ID NO:24) shown in Table 9A. A GPCR9a ORF begins with a Kozak consensus ATG initiation codon at nucleotides 469-471 and ends with a TGA codon at nucleotides 1447-1449. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 9A, and the start and stop codons are in bold letters.

Table 9A. GPCR9 Nucleotide Sequence (SEQ ID NO:24)

TGCTATAGCCCCAGCACTTGATACCTAGCACAGAATAGGTACTTAATAAATACTTAGTGGATGAAT
AAATCTGAAATACTATGGCCATAATTTGGTCACATGAAGCCGTAATGTAGAAAAGATGCTTCCTGT
TAATGACCAAAAACACTTTGGATTCCAAACGATCATTTTAAACATGAATCTTTCTCTGCTGTCTCC
TCTGACCCCATCCTGGGGAGAGCAGAGAGGAGCCTAGGGGACTAGAATGTGCCCCATCCTCCCCTC
AGTGACGTCCACAGAACTGCAGCGCTGAGAAGGCCAGATTGCAGATCTGAAGTCCAACTCCCTCAT

TATACAGATGGTGAAACTAAATTCCAGAGAGGGAGGCTGACCTGCTGCAGCTCAGACATCAGGTCA
CTGGGCTCCCAGGCCAGTTGGAGCTTTTTCCAAAAAGCTGGGTGGTCCAGATGGAAAAGGAGAGAG
AATGAGATGAAGTGGGCAAACCAGACAGCTGTGACGGAATACGTCCTGATGGGGCTACACGAGCAC
TGTAACCTGGAGGTCGTCCTGTTTGTGTTCTGCCTGGGCATCTACTCCGTGAATGTGTTGGGGAAC
GCCCTCCTCATAGGGCTGAACGTGCTGCACCCTCGCCTGCACAACCCCATGTACTTCCTTCTCAGC
AACCTCTCCCTCATGGACATCTGCGGCACCTCCTCCTTTGTGCCTCTCATGCTAGACAATTTCCTG
GAAACCCAGAGGACCATTTCCTTCCCTGGCTGTGCCCTGCAGATGTACCTGACCCTGGCGCTGGGA
TCAACGGAGTGCCTGCTGCTGGCTGTGATGGCATATGACCGTTATGTGGCTATCTGCCAGCCGCTT
AGGTACCCAGAGCTCATGAGTGGGCAGACCTGCATGCAGATGGCAGCGCTGAGCTGGGGGACAGGC
TTTGCCAACTCACTGCTACAGTCCATCCTTGTCTGGCACCTCCCCTTCTGTGGCCACGTCATCAAC
TACTTCTATGAGATCTTGGCAGTGCTAAAACTGGCCTGTGGGGACATCTCCCTCAATGCGCTGGCA
TTAATGGTGGCCACAGCCGTCCTGACACTGGCCCCCCTCTTGCTCATCTGCCTGTCTTACCTTTTC
ATCCTGTCTGCCATCCTTAGGGTACCCTCTGCTGCAGGCCGGTGCAAAGCCTTCTCCACCTGCTCA
GCCCACCGCACAGTGGTGGTGGTTTTTTATGGGACAATCTCCTTCATGTACTTCAAACCCAAGGCC
AAGGATCCCAACGTGGATAAGACTGTCGCATTGTTCTACGGGGTTGTGACGCCCTCGCTGAACCCC
ATCATTTACAGCCTGAGGAATGCAGAGGTGAAAGCTGCCGTCCTAACTCTGCTGAGAGGAGGTTTG
CTCTCCAGGAAAGCATCCCACTGCTACTGCTGCCCTCTGCCCCTGTCAGCTGGCATAGGCTAGGTT
GTGCTGTGGTCATGACCTCAAACCTTGAGAGGCTTAAAGCCATTAAGGTTTGTTTCTTGCTCCTGA
TGCAGGTCCACCAGAGGCTGGTGGGGCTTCTGCTCCGCATCATGGTCTTCACCCCTCTGGGACTCA
GGATGACAAAACAGCTACCATTGGGAACACTGCTGGTCACCATGACAAAAAGAAAAGGGAAAGTAA
CAAAGCCTACACTGACTCTTAAAGCTTCTACTCAGAAGTGGCTGTGTTGCCTCCACCTACATTTCA
GTGGCCAACACAATGGCAACAGGAAGGCACAGGACCACACCTATTGTTAAGGGGGAAAAGCACACT
ATCGTGTGTCTGGATGGCAAACGAGAGGGACAGAGAGATTTGTGAATGGCCTAATGACTACCACAC
CAGCTGACAGTGTCAACCCAAGAGCTATGGGAGGTTTGGCTTTCTTTATCCTGACCATCTATCCTT
CACGGGCTGCTGCCAAGTTAATCGTCCCAAGAAAGCTCTGGTTAGCTCACGTGTGGTAGCTTTATA
CTGAGTCAACCAAACTAGGCTAGAGGGTGTGGGTTAGGGTTGGCCACA

The sequence of GPCR9a was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. The cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some portion thereof.

The cDNA coding for the GPCR9a sequence was cloned by the polymerase chain reaction (PCR). Primers were designed based on *in silico* predictions of the full length or some portion (one or more exons) of the cDNA/protein sequence of the invention. The DNA sequence and protein sequence for a novel Olfactory Receptor-like gene were obtained by exon linking and are reported here as GPCR9a.These primers and methods used to amplify GPCR9 a cDNA are described in the Examples.

The GPCR9a polypeptide (SEQ ID NO:25) encoded by SEQ ID NO:24 is 326 aa in length, has a molecular weight of 35713.69 Daltons, and is presented using the one-letter amino acid code in Table 9B. The Psort profile for GPCR9 predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.600. In alternative embodiments, a GPCR9 polypeptide is located to the Golgi body with a certainty of 0.400, the endoplasmic reticulum (membrane) with a certainty of 0.300, or a mitochondrial

inner membrane with a certainty of 0.300. The Signal P software program predicts no likely signal cleavage site for a GPCR9 peptide.

Table 9B. GPCR9a protein sequence (SEQ ID NO:25)

```
MKWANQTAVTEYVLMGLHEHCNLEVVLFVFCLGIYSVNVLGNALLIGLNVLHPRLHNPMYFLLSNLS
LMDICGTSSFVPLMLDNFLETQRTISFPGCALQMYLTLALGSTECLLLAVMAYDRYVAICQPLRYPE
LMSGQTCMQMAALSWGTGFANSLLQSILVWHLPFCGHVINYFYEILAVLKLACGDISLNALALMVAT
AVLTLAPLLLICLSYLFILSAILRVPSAAGRCKAFSTCSAHRTVVVVFYGTISFMYFKPKAKDPNVD
KTVALFYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGGLLSRKASHCYCCPLPLSAGIG
```

GPCR9b

In an alternative embodiment, a GPCR9 variant is the novel GPCR9b (alternatively referred to herein as 21629632_EXT, Spliced AL133410), which includes the 1069 nucleotide sequence (SEQ ID NO:26) shown in Table 9C. The GPCR9b ORF begins with a Kozak consensus ATG initiation codon at nucleotides 31-33 and ends with a TGA codon at nucleotides 1009-1011, which are in bold letters in Table 9C. Putative untranslated regions upstream from the initiation codon and downstream from the termination codon are underlined in Table 9A, and the start and stop codons are in bold letters.

Table 9C. GPCR9b Nucleotide Sequence (SEQ ID NO:26)

```
TGGTCCAGATGGAAAAGGAGAGAGAATGAGATGAAGTGGGCAAACCAGACAGCTGTGACGGAATACG
TCCTGATGGGGCTACACGAGCACTGTAACCTGGAGGTGGTCCTGTTTGTGTTCTGCCTGGGCATCTA
CTCCGTGAATGTGTTGGGGAACGCCCTCCTCATAGGGCTGAACGTGCTGCACCCTCGCCTGCACAAC
CCCATGTACTTCCTTCTCAGCAACCTCTCCCTCATGGACATCTGCGGCACCTCCTCCTTTGTGCCTC
TCATGCTAGACAATTTCCTGGAAACCCAGAGGACCATTTCCTTCCCTGGCTGTGCCCTGCAGATGTA
CCTGACCCTGGCGCTGGGATCAACGGAGTGCCTGCTGCTGGCTGTGATGGCATATGACCGTTATGTG
GCTATCTGCCAGCCGCTTAGGTACCCAGAGCTCATGAGTGGGCAGACCTGCATGCAGATGGCAGCGC
TGAGCTGGGGGACAGGCTTTGCCAACTCACTGCTACAGTCCATCCTTGTCTGGCACCTCCCCTTCTG
TGGCCACGTCATCAACTACTTCTATGAGATCTTGGCAGTGCTAAAACTGGCCTGTGGGGACATCTCC
CTCAATGCGCTGGCATTAATGGTGGCCACAGCCGTCCTGACACTGGCCCCCCTCTTGCTCATCTGCC
TGTCTTACCTTTTCATCCTGTCTGCCATCCTTAGGGTACCCTCTGCTGCAGGCCGGTGCAAAGCCTT
CTCCACCTGCTCAGCCCACCGCACAGTGGTGGTGGTTTTTTATGGGACAATCTCCTTCATGTACTTC
AAACCCAAGGCCAAGGATCCCAACGTGGATAAGACTGTCGCATTGTTCTACGGGGTTGTGACGCCCT
CGCTGAACCCCATCATTTACAGCCTGAGGAATGCAGAGGTGAAAGCTGCCGTCCTAACTCTGCTGAG
AGGAGGTTTGCTCTCCAGGAAAGCATCCCACTGCTACTGCTGCCCTCTGCCCCTGTCAGCTGGCATA
GGCTAGGTTGTGCTGTGGTCATGACCTCAAACCTTGAGAGGCTTAAAGCCATTAAGGTTTGTTT
```

The GPCR9 protein encoded by SEQ ID NO:26 is identical to SEQ ID NO:25.

GPCR9 Clones

Unless specifically addressed as GPCR9a or GPCR9b, any reference to GPCR9 is assumed to encompass all variants. The GPCR9 nucleic acid sequences differ where GPCR9a extends further in both the 4' and 3' untranslated regions..

In a search of sequence databases, it was found, for example, that the GPCR9b nucleic acid sequence has 856 of 1069 (80%) identical to a *Mouse* Olfactory receptor mRNA (GENBANK-ID: MMU133430). The full GPCR9 amino acid sequence was found to have 231 of 310 amino acid residues (74%) identical to, and 249 of 310 residues (80%) similar to, the 315 amino acid residue protein from *Mouse* (ptnr: SPTREMBL-ACC: Q9QZ17). Additional BLAST results are shown in Table 9E.

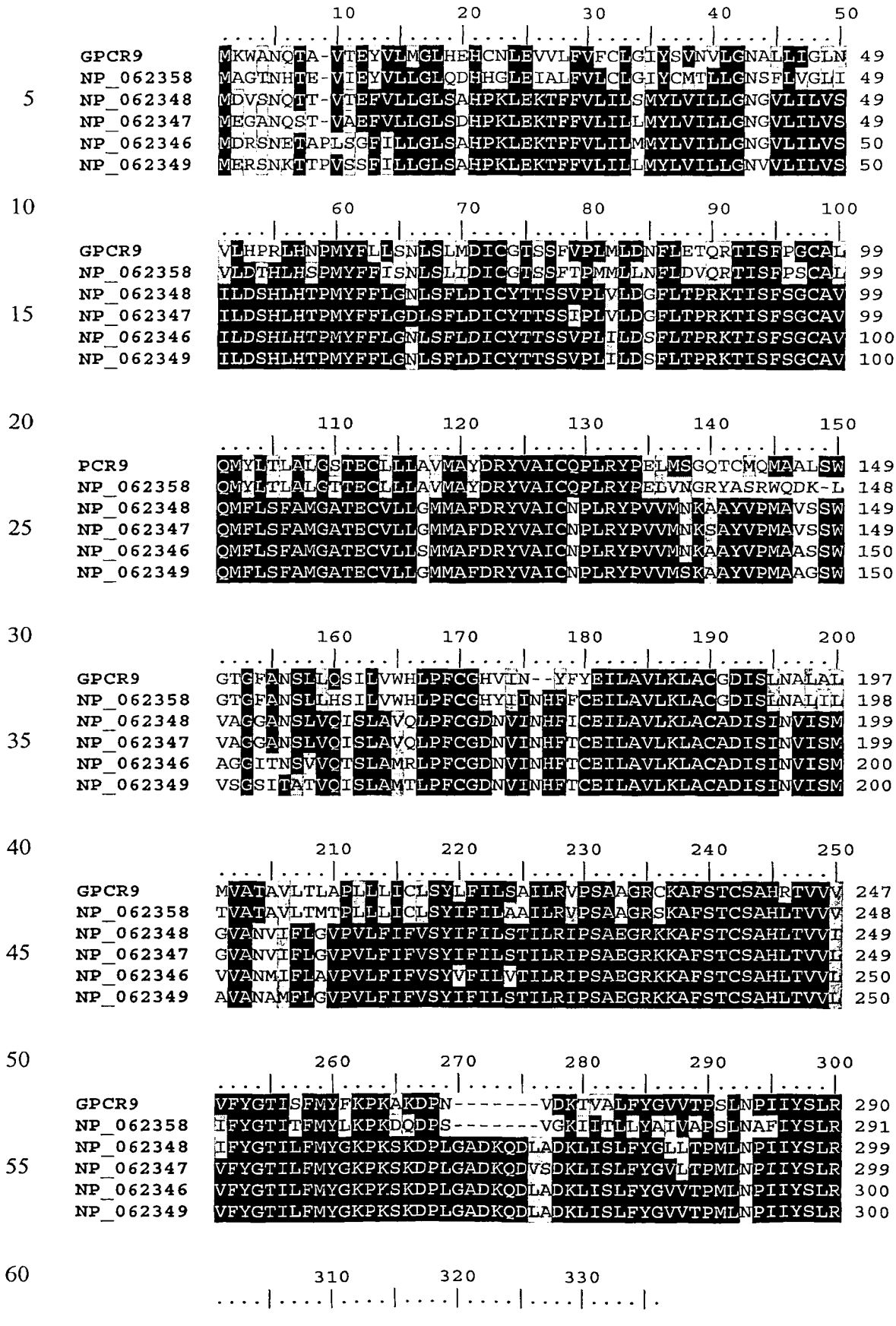
Table 9E. BLAST results for GPCR9

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|11464983|ref|NP_062358.1|; gi|5869927|emb|CAB55598.1 (AJ133430); gi|8919698|emb|CAB96153.1 | (AJ251155)	olfactory receptor 70 [Mus musculus]	315	206/315 (65%)	234/315 (73%)	5e-99
gi|11276079|ref|NP_062348.1|; gi|5869920|emb|CAB55594.1 | (AJ133426); gi|8919693|emb|CAB96148.1 | (AJ251154)	olfactory receptor 37c [Mus musculus]	318	170/310 (54%)	207/310 (65%)	2e-77
gi|11276077|ref|NP_06; 2347.1|; gi|5869918|emb|CAB55593.1 | (AJ133425); gi|8919694|emb|CAB96149.1 | (AJ251154)	olfactory receptor 37b [Mus musculus]	318	169/311 (54%)	209/311 (66%)	2e-77
gi|11276075|ref|NP_062346.1|; gi|5869916|emb|CAB55592.1 | (AJ133424); gi|8919692|emb|CAB96147.1 | (AJ251154)	olfactory receptor 37a [Mus musculus]	319	167/312 (53%)	211/312 (67%)	5e-76
gi|11464981|ref|NP_062349.1|; gi|5869923|emb|CAB55596.1 | (AJ133428); gi|8919695|emb|CAB96151.1 | (AJ251154)	olfactory receptor 37e [Mus musculus]	319	165/305 (54%)	203/305 (66%)	3e-74

A multiple sequence alignment is given in Table 9F, with the GPCR9 protein of the invention being shown on line 1, in a ClustalW analysis comparing GPCR9 with related protein sequences, shown in Table 9E.

Table 9F. Information for the ClustalW proteins:

1. GPCR9; SEQ ID NO:25
2. gi|11464983|ref|NP_062358.1| olfactory receptor 70 [Mus musculus]; SEQ ID NO:70
3. gi|11276079|ref|NP_062348.1| olfactory receptor 37c [Mus musculus]; SEQ ID NO:71
4. gi|11276077|ref|NP_062347.1| olfactory receptor 37b [Mus musculus]; SEQ ID NO:72
5. gi|11276075|ref|NP_062346.1| olfactory receptor 37a [Mus musculus]; SEQ ID NO:73
6. gi|11464981|ref|NP_062349.1| olfactory receptor 37e [Mus musculus]; SEQ ID NO:74


10 20 30 40 50
GPCR9 MKWANQTA-VTEYVLMGLHEHCNLEVVLFVFCLGIYSVNVLGNALLIGLN 49
NP_062358 MAGTNHTE-VIEYVLLGLQDHHGLEIALFVLCLGIYCMTLLGNSFLVGLI 49
NP_062348 MDVSNQTT-VTEFVLLGLSAHPKLEKTFFVLILSMYLVILLGNGVLILVS 49
NP_062347 MEGANQST-VAEFVLLGLSDHPKLEKTFFVLILLMYLVILLGNGVLILVS 49
NP_062346 MDRSNETAPLSGFILLGLSAHPKLEKTFFVLILMMYLVILLGNGVLILVS 50
NP_062349 MERSNKTTPVSSFILLGLSAHPKLEKTFFVLILLMYLVILLGNVVLILVS 50
60 70 80 90 100
GPCR9 VLHPRLHNPMYFLLSNLSLMDICGTSSFVPLMLDNFLETQRTISFPGCAL 99
NP_062358 VLDTHLHSPMYFFISNLSLIDICGTSSFTPMMLLNFLDVQRTISFPSCAL 99
NP_062348 ILDSHLHTPMYFFLGNLSFLDICYTTSSVPLVLDGFLTPRKTISFSGCAV 99
NP_062347 ILDSHLHTPMYFFLGDLSFLDICYTTSSIPLVLDGFLTPRKTISFSGCAV 99
NP_062346 ILDSHLHTPMYFFLGNLSFLDICYTTSSVPLILDSFLTPRKTISFSGCAV 100
NP_062349 ILDSHLHTPMYFFLGNLSFLDICYTTSSVPLILDSFLTPRKTISFSGCAV 100
110 120 130 140 150
PCR9 QMYLTLALGSTECLLLAVMAYDRYVAICQPLRYPELMSGQTCMQMAALSW 149
NP_062358 QMYLTLALGTTECLLLAVMAYDRYVAICQPLRYPELVNGRYASRWQDK-L 148
NP_062348 QMFLSFAMGATECVLLGMMAFDRYVAICNPLRYPVVMNKAAYVPMAVSSW 149
NP_062347 QMFLSFAMGATECVLLGMMAFDRYVAICNPLRYPVVMNKSAYVPMAVSSW 149
NP_062346 QMFLSFAMGATECVLLSMMAFDRYVAICNPLRYPVVMNKAAYVPMAASSW 150
NP_062349 QMFLSFAMGATECVLLGMMAFDRYVAICNPLRYPVVMSKAAYVPMAAGSW 150
160 170 180 190 200
GPCR9 GTGFANSLLQSILVWHLPFCGHVIN--YFYEILAVLKLACGDISLNALAL 197
NP_062358 GTGFANSLLHSILVWHLPFCGHYIINHFFCEILAVLKLACGDISLNALIL 198
NP_062348 VAGGANSLVQISLAVQLPFCGDNVINHFICEILAVLKLACADISINVISM 199
NP_062347 VAGGANSLVQISLAVQLPFCGDNVINHFTCEILAVLKLACADISINVISM 199
NP_062346 AGGITNSVVQTSLAMRLPFCGDNVINHFTCEILAVLKLACADISINVISM 200
NP_062349 VSGSITATVQISLAMTLPFCGDNVINHFTCEILAVLKLACADISINVISM 200
210 220 230 240 250
GPCR9 MVATAVLTLAPLLLICLSYLFILSAILRVPSAAGRCKAFSTCSAHRTVVV 247
NP_062358 TVATAVLTMTPLLLICLSYIFILAAILRVPSAAGRSKAFSTCSAHLTVVV 248
NP_062348 GVANVIFLGVPVLFIFVSYIFILSTILRIPSAEGRKKAFSTCSAHLTVVI 249
NP_062347 GVANVIFLGVPVLFIFVSYIFILSTILRIPSAEGRKKAFSTCSAHLTVVI 249
NP_062346 VVANMIFLAVPVLFIFVSYVFILVTILRIPSAEGRKKAFSTCSAHLTVVI 250
NP_062349 AVANAMFLGVPVLFIFVSYIFILSTILRIPSAEGRKKAFSTCSAHLTVVI 250
260 270 280 290 300
GPCR9 VFYGTISFMYFKPKAKDPN-------VDKIVALFYGVVTPSLNPIIYSLR 290
NP_062358 IFYGTITFMYLKPKDQDPS-------VGKIITLLYAIVAPSLNAFIYSLR 291
NP_062348 IFYGTILFMYGKPKSKDPLGADKQDLADKLISLFYGLLTPMLNPIIYSLR 299
NP_062347 VFYGTILFMYGKPKSKDPLGADKQDVSDKLISLFYGVLTPMLNPIIYSLR 299
NP_062346 VFYGTILFMYGKPKSKDPLGADKQDLADKLISLFYGVVTPMLNPIIYSLR 300
NP_062349 VFYGTILFMYGKPKSKDPLGADKQDLADKLISLFYGVVTPMLNPIIYSLR 300
310 320 330

```
GPCR9       NAEVKAAVLTLLRGGLLSRKASHCYCCPLPLSAGIG 326
NP_062358   NSEVKAAVTALLWGGLLTRKMSHF------------ 315
NP_062348   NKDVKAAVRNLASHRCLTF----------------- 318
NP_062347   NKDVKAAVRNLVGQKCLIQ----------------- 318
NP_062346   NKDVRAAVRNLVGQKHLTE----------------- 319
NP_062349   NKDVKAAVTNLVGQKHFKW----------------- 319
```

DOMAIN results for GPCR9 were collected from the Conserved Domain Database (CDD) with Reverse Position Specific BLAST. This BLAST samples domains found in the Smart and Pfam collections. The results are listed in Table 9G with the statistics and domain description. The 7tm_1, a seven transmembrane receptor (rhodopsin family), was shown to have significant homology to GPCR9. An alignment of GPCR9 residues 41-287 (SEQ ID NO:26) with 7tm_1 residues 1-254 (SEQ ID NO:34) are shown in Table 9G.

Table 9G. DOMAIN results for GPCR9

PSSMs producing significant alignments:	Score (bits)	E value
gnl|Pfam|pfam00001 7tm_1, 7 transmembrane receptor (rhodopsin family)	88.2	6e-19

```
                    10        20        30        40        50
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               GNALLIGLNVLHPRLHNPMYFLLSNLSLMDICGTSSFVPLMLDNFLETQR
Pfam|pfam00001      GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPFWALYYLVGGDW

                    60        70        80        90        100
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               TISFPGCALQMYLTLALGSTECLLLAVMAYDRYVAICQPLRYPELMSGQT
Pfam|pfam00001      VFGDALCKLVGALFWVNGYASILLLTAISIDRYLAIVHPLRYRRIRTPRR

                    110       120       130       140       150
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               CMQMAALSWGTGFANSLLQSILVW-------HL----PFCGHVINYFYEI
Pfam|pfam00001      AKVLILLVWVLALLLSLPPLLFSWLRTVEEGNT----TVCLIDFPEESVK

                    160       170       180       190       200
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               LAVLKLACGDISLNALALMVATAVLTLAPLLLICLSYLFILS--------
Pfam|pfam00001      RSYVLLSTLVG-----FVLPLLVILVCYTRILRTLRKRARSQ--------

                    210       220       230       240       250
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               ----------------------------------------------------
Pfam|pfam00001      ----------------------------------------------------

                    260       270       280       290       300
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               ----------------------------------------AILRVPS
Pfam|pfam00001      ----------------------------------------RSLKRRS

                    310       320       330       340       350
                    |....|....|....|....|....|....|....|....|....|....|
GPCR9               AAGRCKAFSTCSAHRTVVVVFYGTISFMYFKPKAKD-----PNVDKTVAL
Pfam|pfam00001      SSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIWRVLPTALLITL
```

```
                          360
                    ....|....|....
GPCR9               FYGVVTPSLNPIIY
Pfam|pfam00001      WLAYVNSCLNPIIY
```

The GPCR disclosed in this invention is expressed in at least the following tissues: Prostate, ovary.

The nucleic acids and proteins of GPCR2 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further above.

The novel nucleic acid encoding the GPCR-like protein of the invention, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR2 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR2 epitope is from about amino acids 5 to 20. In other specific embodiments, GPCR2 epitopes are from about amino acids 125 to 145, from about amino acids 230 to 240 and from about amino acids 255 to 275.

GPCR10

The disclosed novel GPCR10 nucleic acid (SEQ ID NO:27) of 1147 nucleotides (also referred to as 1823044_EXT) encoding a novel olfactory receptor-like protein is shown in Table 10A. The GPCR disclosed in this invention maps to chromosome 1. An ORF begins with an ATG initiation codon at nucleotides 17-19 and ends with a TAG codon at nucleotides 1061-1063. A putative untranslated region upstream from the initiation codon and downstream from the termination codon is underlined in Table 10A, and the start and stop codons are in bold letters.

Table 10A. GPCR10 Nucleotide Sequence (SEQ ID NO:27)

GCCTAGGTGAAACCTC**ATG**GACAACATCACCTGGATGGCCAGCCACACTGGATGGTCGGAT
TTCATCCTGATGGGACTCTTCAGACAATCCAAACATCCAATGGCCAATATCACCTGGATGG
CCAACCACACTGGATGGTCGGATTTCATCCTGTTGGGACTCTTCAGACAATCCAAACATCC
AGCACTACTTTGTGTGGTCATTTTTGTGGTTTTCCTGATGGCGTTGTCTGGAAATGCTGTC
CTGATCCTTCTGATACACTGTGACGCCCACCTCCACACCCCCATGTACTTTTTCATCAGTC
AATTGTCTCTCATGGACATGGCGTACATTTCTGTCACTGTGCCCAAGATGCTCCTGGACCA
GGTCATGGGTGTGAATAAGATCTCAGCCCCTGAGTGTGGGATGCAGATGTTCTTCTACGTG

ACACTAGCAGGTTCAGAATTTTTCCTTCTAGCCACCATGGCCTATGACCGCTACGTGGCCA
TCTGCCATCCTCTCCGTTACCCTGTCCTCATGAACCATAGGGTGTGTCTCTTCCTGTCATC
AGGCTGCTGGTTCCTGGGCTCAGTGGATGGCTTCACATTCACTCCCATCACCATGACCTTC
CCCTTCCGTGGATCCCGGGAGATTCATCATTTCTTCTGTGAAGTTCCTGCTGTATTGAATC
TCTCCTGCTCAGACACCTCACTCTATGAGATTTTCATGTACTTGTGCTGTGTCCTCATGCT
CCTCATCCCTGTGGTGATCATTTCAAGCTCCTATTTACTCATCCTCCTCACCATCCACGGG
ATGAACTCAGCAGAGGGCCGGAAAAAGGCCTTTGCCACCTGCTCCTCCCACCTGACTGTGG
TCATCCTCTTCTATGGGGCTGCCATCTACACCTACATGCTCCCCAGCTCCTACCACACCCC
TGAGAAGGACATGATGGTATCTGTCTTCTATACCATCCTCACTCCAGTGGTGAACCCTTTA
ATCTATAGTCTTAGGAATAAGGATGTCATGGGGGCTCTGAAGAAAATGTTAACAGTGGAAC
CTGCCTTTCAAAAAGCTATGGAG**TAG**ACCATTTTGAGAGTAATTTACTTTTCCTTCTCTCT
GCACTTCACATATGAGAATGTTATACCAGTGTTATTTCCCAGACTCCAA

The GPCR10 protein (SEQ ID NO:28) encoded by SEQ ID NO:27 has 348 amino acid residues and is presented using the one-letter code in Table 10B. The predicted molecular weight of GPCR10 protein is approximately 39411.93 Daltons. The Psort profile for GPCR10 predicts that this sequence has a signal peptide and is likely to be localized at the plasma membrane with a certainty of 0.6. In alternative embodiments, GPCR10 is located in the golgi body with a certainty of 0.4, the endoplasmic reticulum (membrane) with a certainty of 0.3 or microbodies (peroxisomes) with a certainty of 0.3. The Signal P predicts a likely cleavage site between positions 19 and 20, i.e., at the dash in the sequence ILM-GL. The protein predicted here is similar to the "Olfactory Receptor-Like Protein Family", some members of which end up localized at the cell surface where they exhibit activity. Therefore, it is likely that this novel olfactory receptor-like proteins is available at the appropriate sub-cellular localization and hence accessible for the therapeutic uses described in this application.

Table 10B. Encoded GPCR10 protein sequence (SEQ ID NO:28)

MDNITWMASHTGWSDFILMGLFRQSKHPMANITWMANHTGWSDFILLGLFRQSKHPALLC
VVIFVVFLMALSGNAVLILLIHCDAHLHTPMYFFISQLSLMDMAYISVTVPKMLLDQVMG
VNKISAPECGMQMFFYVTLAGSEFFLLATMAYDRYVAICHPLRYPVLMNHRVCLFLSSGC
WFLGSVDGFTFTPITMTFPFRGSREIHHFFCEVPAVLNLSCSDTSLYEIFMYLCCVLMLL
IPVVIISSSYLLILLTIHGMNSAEGRKKAFATCSSHLTVVILFYGAAIYTYMLPSSYHTP
EKDMMVSVFYTILTPVVNPLIYSLRNKDVMGALKKMLTVEPAFQKAME

In a search of sequence databases, it was found, for example, that the nucleic acid sequence of this invention has 434 of 488 bases (88%) identical to a *gibbon* olfactory receptor mRNA (GENBANK-ID: AF179779). The full amino acid sequence of the protein of the invention was found to have 147 of 223 amino acid residues (65%) identical to, and 177 of 223 residues (79%) similar to the 223 amino acid residue protein from *mouse*

(ptnr:SPTREMBL-ACC: Q62342). GPCR10 also has homology to the proteins shown in the BLASTP data in Table 10C.

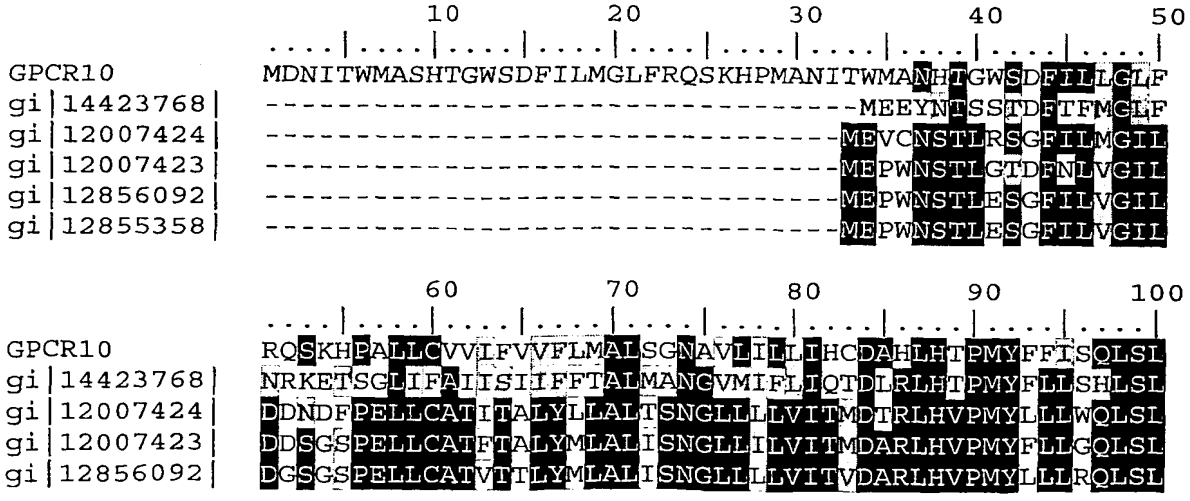
Table 10C. BLAST results for GPCR10

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|14423768|sp|O43869	O2T1_HUMAN OLFACTORY RECEPTOR 2T1 (OLFACTORY RECEPTOR 1-25) (OR1-25)	311	165/300 (55%)	202/300 (67%)	2e-83
gi|12007424|gb|AAG45197.1| (AF321234)	T3 olfactory receptor [Mus musculus]	315	139/303 (45%)	186/303 (60%)	3e-68
gi|12007423|gb|AAG45196.1| (AF321234)	T2 olfactory receptor [Mus musculus]	316	138/301 (45%)	184/301 (60%)	8e-68
gi|12856092|dbj|BAB30564.1| (AK017036)	putative [Mus musculus]	316	134/301 (44%)	186/301 (61%)	2e-66
gi|12855358|dbj|BAB30304.1| (AK016560)	putative [Mus musculus]	316	134/301 (44%)	186/301 (61%)	2e-66

A multiple sequence alignment is given in Table 10D, with the protein of the invention 1823044_EXT being shown on lines 1 in Table 10D in a ClustalW analysis comparing the protein of the invention with related protein sequences. This BLASTP data is displayed graphically in the Clustal W in Table 10D.

Table 10D. ClustalW Analysis of GPCR10

1) GPCR10; SEQ ID NO:28
2) >gi|14423768|sp|O43869|O2T1_ Human Olfactory Receptor 2T1 (OR1-25) ; SEQ ID NO:75
3) >gi|12007424|gb|AAG45197.1| (AF321234) T3 olfactory receptor [Mus musculus]; SEQ ID NO:76
4) >gi|12007423|gb|AAG45196.1| (AF321234) T2 olfactory receptor [Mus musculus]; SEQ ID NO:77
5) >gi|12856092|dbj|BAB30564.1| (AK017036) putative [Mus musculus]; SEQ ID NO:78
6) >gi|12855358|dbj|BAB30304.1| (AK016560) putative [Mus musculus]; SEQ ID NO:79



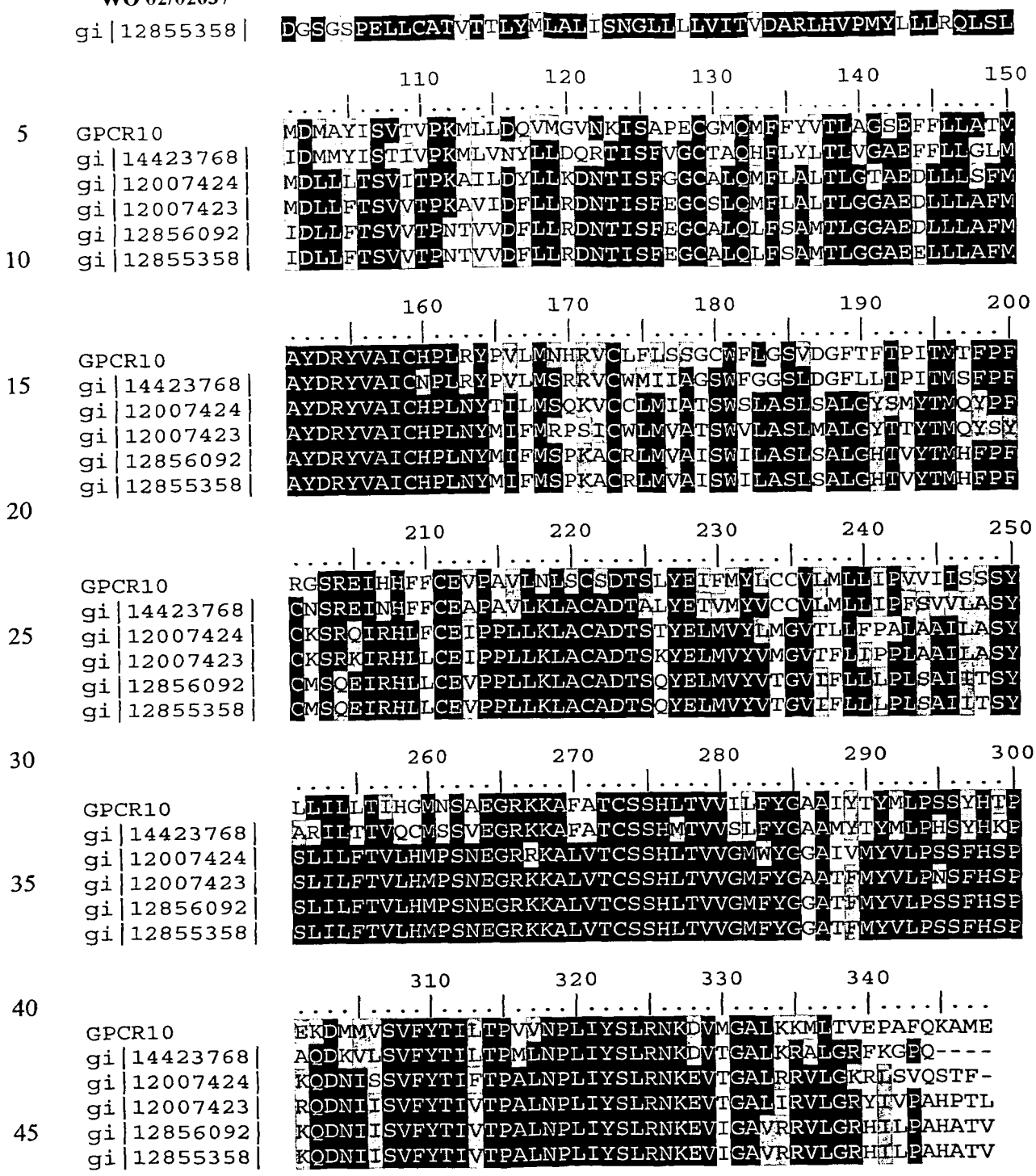


Table 10E lists the domain description from DOMAIN analysis results against GPCR10. This indicates that the GPCR10 sequence has properties similar to those of other proteins known to contain this domain as well as to the 254 amino acid 7tm domain (SEQ ID NO:34) itself.

Table 10E Domain Analysis of GPCR10

```
PSSMs producing significant alignments:                                Score E
                                                                       (bits) value
gnl|Pfam|pfam00001  7tm_1, 7 transmembrane receptor (rhodopsin family) 90.9  1e-19

                          10        20        30        40        50
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    GNAVLILLIHCDAHLHTPMYFFISQLSLMDMAYISVTVPKMLLDQVMG--
Pfam|pfam00001            GNLLVILVILRTKKLRTPTNIFLLNLAVADLLFLLTLPFWALYYLVGG--

                          60        70        80        90        100
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    ----VNKISAPECGMQMFFYVTLAGSEFFLLATMAYDRYVAICHPLRYPV
Pfam|pfam00001            ----DWVFGDALCKLVGALFVVNGYASILLLTAISIDRYLAIVHPLRYRR

                          110       120       130       140       150
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    LMNHRVCLFLSSGCWFLGSVDGFTFTPITMTFPFRGSRE----IHHFFCE
Pfam|pfam00001            IRTPRRAKVLILLVWVLALLLSLPPLLFSWLRTVEEGNT----TVCLIDF

                          160       170       180       190       200
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    VPAVLNLSCSDTSLYEIFMYLCCVLMLLIPVVIISSSYLLILL-------
Pfam|pfam00001            PEESVKRSYVLLSTLVGFVLPLLVILVCYTRILRTLRKRARSQ-------

                          210       220       230       240       250
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    ----------------------------------------------------
Pfam|pfam00001            ----------------------------------------------------

                          260       270       280       290       300
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    ----------------------------------------------TIHGMN
Pfam|pfam00001            ----------------------------------------------RSLKRR

                          310       320       330       340       350
                          |....|....|....|....|....|....|....|....|....|....|
GPCR10                    SAEGRKKAFATCSSHLTVVLLFYGAAIYTYMLP----SSY-HTPEKDMMV
Pfam|pfam00001            SSSERKAAKMLLVVVVVFVLCWLPYHIVLLLDSLCLLSIW-RVLPTALLI

                          360
                          |....|....|....|.
GPCR10                    SVFYTILTPVVNPLIY
Pfam|pfam00001            TLWLAYVNSCLNPIIY
```

The olfactory receptor disclosed in this invention is expressed in at least the following tissues: lymph node, ovary.

The nucleic acids and proteins of GPCR10 are useful in potential therapeutic applications implicated in various GPCR-related pathological disorders and/or OR-related pathological disorders, described further herein. The novel GPCR10 nucleic acid, or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed.

These materials are further useful in the generation of antibodies that bind immunospecifically to the novel substances of the invention for use in therapeutic or

diagnostic methods. These antibodies may be generated according to methods known in the art, using prediction from hydrophobicity charts, as described in the "Anti-GPCRX Antibodies" section below. The disclosed GPCR10 protein has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated GPCR10 epitope is from about amino acids 10 to 18. In additional embodiments, GPCR10 epitopes are from about amino acids 20 to 30, from about amino acids 42 to 50, from about amino acids 190 to 210, from about amino acids 260 to 270, from about amino acids 280 to 308 and from about amino acids 325 to 340.

A summary of the GPCRX nucleic acids and proteins of the invention is provided in Table 11.

TABLE 11: Summary Of Nucleic Acids And Proteins Of The Invention

Name	Tables	Clone; Description of Homolog	Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO
GPCR1	1A, 1B	GPCR1a: CG54326_02	1	2
	1C, 1D	GPCR1b: AP001804_A	3	4
GPCR2	2A, 2B	GPCR2a: CG54335_02	5	6
	2C, 2D	GPCR2b: AP001804_B	7	8
GPCR3	3A, 3B	GPCR3: AP001804_C	9	10
GPCR4	4A, 4B	GPCR4: AP001804_D	11	12
GPCR5	5A, 5B,	GPCR5a: CG56040_01	13	14
	5C, 5D	GPCR5b: AP001804_E	15	16
GPCR6	6A, 6B	GPCR6a: CG56025-01	17	18
	6C	GPCR6b: AP001804_B	19	
GPCR7	7A, 7B	GPCR7: 20722608_EXT	20	21
GPCR8	8A, 8B	GPCR8: CG-SC931712	22	23
GPCR9	9A, 9B	GPCR9a: 21629632.0.20	24	25
	9C	GPCR9b: 21629632_E	26	
GPCR10	10A, 10B	GPCR10: 1823044_EXT	27	28

GPCRX Nucleic Acids and Polypeptides

One aspect of the invention pertains to isolated nucleic acid molecules that encode GPCRX polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify GPCRX-encoding nucleic acids (*e.g.*, GPCRX mRNAs) and fragments for use as PCR primers for the amplification and/or mutation of GPCRX nucleic acid molecules. As used herein, the term

"nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is comprised double-stranded DNA.

An GPCRX nucleic acid can encode a mature GPCRX polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full-length gene product, encoded by the corresponding gene. Alternatively, it may be defined as the polypeptide, precursor or proprotein encoded by an ORF described herein. The product "mature" form arises, again by way of nonlimiting example, as a result of one or more naturally occurring processing steps as they may take place within the cell, or host cell, in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the cleavage of the N-terminal methionine residue encoded by the initiation codon of an ORF, or the proteolytic cleavage of a signal peptide or leader sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation, myristoylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

The term "probes", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as approximately, *e.g.*, 6,000 nt, depending upon the specific use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as utilized herein, is one, which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated GPCRX nucleic acid molecules can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (*e.g.*, brain, heart, liver, spleen, etc.). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or a complement of this aforementioned nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 as a hybridization probe, GPCRX molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to GPCRX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or

100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or a complement thereof. Oligonucleotides may be chemically synthesized and may also be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or a portion of this nucleotide sequence (*e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of an GPCRX polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 is one that is sufficiently complementary to the nucleotide sequence shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a

similar or opposite metabolic activity compared to wild type. Homologs are nucleic acid sequences or amino acid sequences of a particular gene that are derived from different species.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of GPCRX polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention, homologous nucleotide sequences include nucleotide sequences encoding for an GPCRX polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, *e.g.*, frog, mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human GPCRX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, as well as a polypeptide possessing GPCRX biological activity. Various biological activities of the GPCRX proteins are described below.

As used herein, "identical" residues correspond to those residues in a comparison between two sequences where the equivalent nucleotide base or amino acid residue in an alignment of two sequences is the same residue. Residues are alternatively described as "similar" or "positive" when the comparisons between two sequences in an alignment show

that residues in an equivalent position in a comparison are either the same amino acid or a conserved amino acid as defined below.

An GPCRX polypeptide is encoded by the open reading frame ("ORF") of an GPCRX nucleic acid. An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon. An ORF that represents the coding sequence for a full protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF may be any part of a coding sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, *e.g.,* a stretch of DNA that would encode a protein of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the human GPCRX genes allows for the generation of probes and primers designed for use in identifying and/or cloning GPCRX homologues in other cell types, *e.g.* from other tissues, as well as GPCRX homologues from other vertebrates. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27; or an anti-sense strand nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27; or of a naturally occurring mutant of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27.

Probes based on the human GPCRX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express an GPCRX protein, such as by measuring a level of an GPCRX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting GPCRX mRNA levels or determining whether a genomic GPCRX gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of an GPCRX polypeptide" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically-active portion of GPCRX" can be prepared by isolating a portion SEQ ID NOS: 1, 3, 5, 7, 9,

11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 that encodes a polypeptide having an GPCRX biological activity (the biological activities of the GPCRX proteins are described below), expressing the encoded portion of GPCRX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of GPCRX.

GPCRX Nucleic Acid and Polypeptide Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 due to degeneracy of the genetic code and thus encode the same GPCRX proteins as that encoded by the nucleotide sequences shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28.

In addition to the human GPCRX nucleotide sequences shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the GPCRX polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the GPCRX genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame (ORF) encoding an GPCRX protein, preferably a vertebrate GPCRX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the GPCRX genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the GPCRX polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the GPCRX polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding GPCRX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the GPCRX cDNAs of the invention can be isolated based on their homology to the human GPCRX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding GPCRX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM

EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well-known within the art. See, *e.g.*, Ausubel, et *al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.*, Ausubel, *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981. *Proc Natl Acad Sci USA* **78**: 6789-6792.

Conservative Mutations

In addition to naturally-occurring allelic variants of GPCRX sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 thereby leading to changes in the amino acid sequences of the encoded

GPCRX proteins, without altering the functional ability of said GPCRX proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the GPCRX proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the GPCRX proteins of the invention are predicted to be particularly non-amenable to alteration. Amino acids for which conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding GPCRX proteins that contain changes in amino acid residues that are not essential for activity. Such GPCRX proteins differ in amino acid sequence from SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequences of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28; more preferably at least about 70% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28; still more preferably at least about 80% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28; even more preferably at least about 90% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28; and most preferably at least about 95% homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28.

An isolated nucleic acid molecule encoding an GPCRX protein homologous to the protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side

chain. Families of amino acid residues having similar side chains have been defined within the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the GPCRX protein is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an GPCRX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for GPCRX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

The relatedness of amino acid families may also be determined based on side chain interactions. Substituted amino acids may be fully conserved "strong" residues or fully conserved "weak" residues. The "strong" group of conserved amino acid residues may be any one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, wherein the single letter amino acid codes are grouped by those amino acids that may be substituted for each other. Likewise, the "weak" group of conserved residues may be any one of the following: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, VLIM, HFY, wherein the letters within each group represent the single letter amino acid code.

In one embodiment, a mutant GPCRX protein can be assayed for (*i*) the ability to form protein:protein interactions with other GPCRX proteins, other cell-surface proteins, or biologically-active portions thereof, (*ii*) complex formation between a mutant GPCRX protein and an GPCRX ligand; or (*iii*) the ability of a mutant GPCRX protein to bind to an intracellular target protein or biologically-active portion thereof; (*e.g.* avidin proteins).

In yet another embodiment, a mutant GPCRX protein can be assayed for the ability to regulate a specific biological function (*e.g.*, regulation of insulin release).

Antisense Nucleic Acids

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or

fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (*e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire GPCRX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of an GPCRX protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, or antisense nucleic acids complementary to an GPCRX nucleic acid sequence of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an GPCRX protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding the GPCRX protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the GPCRX protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of GPCRX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of GPCRX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of GPCRX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (*e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine,

xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an GPCRX protein to thereby inhibit expression of the protein (*e.g.*, by inhibiting transcription and/or translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (*e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient nucleic acid molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other. See, *e.g.*, Gaultier, *et al.*, 1987. *Nucl. Acids Res.* **15**:

6625-6641. The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (see, *e.g.*, Inoue, *et al.* 1987. *Nucl. Acids Res.* **15**: 6131-6148) or a chimeric RNA-DNA analogue (see, *e.g.*, Inoue, *et al.*, 1987. *FEBS Lett.* **215**: 327-330.

Ribozymes and PNA Moieties

Nucleic acid modifications include, by way of non-limiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes as described in Haselhoff and Gerlach 1988. *Nature* 334: 585-591) can be used to catalytically cleave GPCRX mRNA transcripts to thereby inhibit translation of GPCRX mRNA. A ribozyme having specificity for an GPCRX-encoding nucleic acid can be designed based upon the nucleotide sequence of an GPCRX cDNA disclosed herein (*i.e.*, SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an GPCRX-encoding mRNA. See, *e.g.*, U.S. Patent 4,987,071 to Cech, *et al.* and U.S. Patent 5,116,742 to Cech, *et al.* GPCRX mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, GPCRX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the GPCRX nucleic acid (*e.g.*, the GPCRX promoter and/or enhancers) to form triple helical structures that prevent transcription of the GPCRX gene in target cells. *See, e.g.,* Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, *et al.* 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the GPCRX nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids. *See, e.g.,* Hyrup, *et al.*, 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (*e.g.*, DNA mimics) in which the deoxyribose phosphate backbone is replaced by

a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, *et al.*, 1996. *supra*; Perry-O'Keefe, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

PNAs of GPCRX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of GPCRX can also be used, for example, in the analysis of single base pair mutations in a gene (*e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S_1 nucleases (*see,* Hyrup, *et al.,* 1996.*supra*); or as probes or primers for DNA sequence and hybridization (*see,* Hyrup, *et al.*, 1996, *supra*; Perry-O'Keefe, *et al.,* 1996. *supra*).

In another embodiment, PNAs of GPCRX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of GPCRX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (*e.g.*, RNase H and DNA polymerases) to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (*see,* Hyrup, et al., 1996. *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup, *et al.,* 1996. *supra* and Finn, *et al.*, 1996. *Nucl Acids Res* 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. *See, e.g.,* Mag, *et al.*, 1989. *Nucl Acid Res* 17: 5973-5988. PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. *See, e.g.,* Finn, *et al.*, 1996. *supra*. Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. *See, e.g.,* Petersen, *et al.*, 1975. *Bioorg. Med. Chem. Lett.* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across

the cell membrane (*see, e.g.*, Letsinger, *et al.*, 1989. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6553-6556; Lemaitre, *et al.*, 1987. *Proc. Natl. Acad. Sci.* 84: 648-652; PCT Publication No. WO88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (*see, e.g.*, Krol, *et al.*, 1988. *BioTechniques* 6:958-976) or intercalating agents (*see, e.g.*, Zon, 1988. *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.

GPCRX Polypeptides

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of GPCRX polypeptides whose sequences are provided in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residues shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28 while still encoding a protein that maintains its GPCRX activities and physiological functions, or a functional fragment thereof.

In general, an GPCRX variant that preserves GPCRX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

One aspect of the invention pertains to isolated GPCRX proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-GPCRX antibodies. In one embodiment, native GPCRX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, GPCRX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, an GPCRX protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the GPCRX protein is derived, or substantially free from chemical

precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of GPCRX proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In one embodiment, the language "substantially free of cellular material" includes preparations of GPCRX proteins having less than about 30% (by dry weight) of non-GPCRX proteins (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-GPCRX proteins, still more preferably less than about 10% of non-GPCRX proteins, and most preferably less than about 5% of non-GPCRX proteins. When the GPCRX protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the GPCRX protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of GPCRX proteins in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of GPCRX proteins having less than about 30% (by dry weight) of chemical precursors or non-GPCRX chemicals, more preferably less than about 20% chemical precursors or non-GPCRX chemicals, still more preferably less than about 10% chemical precursors or non-GPCRX chemicals, and most preferably less than about 5% chemical precursors or non-GPCRX chemicals.

Biologically-active portions of GPCRX proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of the GPCRX proteins (*e.g.*, the amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28) that include fewer amino acids than the full-length GPCRX proteins, and exhibit at least one activity of an GPCRX protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the GPCRX protein. A biologically-active portion of an GPCRX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acid residues in length.

Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native GPCRX protein.

In an embodiment, the GPCRX protein has an amino acid sequence shown in SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28. In other embodiments, the GPCRX

protein is substantially homologous to SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, and retains the functional activity of the protein of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below. Accordingly, in another embodiment, the GPCRX protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, and retains the functional activity of the GPCRX proteins of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28.

Determining Homology Between Two or More Sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. *See,* Needleman and Wunsch, 1970. *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the

number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

Chimeric and Fusion Proteins

The invention also provides GPCRX chimeric or fusion proteins. As used herein, an GPCRX "chimeric protein" or "fusion protein" comprises an GPCRX polypeptide operatively-linked to a non-GPCRX polypeptide. An "GPCRX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an GPCRX protein (SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28), whereas a "non-GPCRX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the GPCRX protein, *e.g.*, a protein that is different from the GPCRX protein and that is derived from the same or a different organism. Within an GPCRX fusion protein the GPCRX polypeptide can correspond to all or a portion of an GPCRX protein. In one embodiment, an GPCRX fusion protein comprises at least one biologically-active portion of an GPCRX protein. In another embodiment, an GPCRX fusion protein comprises at least two biologically-active portions of an GPCRX protein. In yet another embodiment, an GPCRX fusion protein comprises at least three biologically-active portions of an GPCRX protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the GPCRX polypeptide and the non-GPCRX polypeptide are fused in-frame with one another. The non-GPCRX polypeptide can be fused to the N-terminus or C-terminus of the GPCRX polypeptide.

In one embodiment, the fusion protein is a GST-GPCRX fusion protein in which the GPCRX sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant GPCRX polypeptides.

In another embodiment, the fusion protein is an GPCRX protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of GPCRX can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is an GPCRX-immunoglobulin fusion protein in which the GPCRX sequences are fused to sequences derived from a member of the immunoglobulin protein family. The GPCRX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between an GPCRX ligand and an GPCRX protein on the surface of a cell, to thereby suppress GPCRX-mediated signal transduction *in vivo*. The GPCRX-immunoglobulin fusion proteins can be used to affect the bioavailability of an GPCRX cognate ligand. Inhibition of the GPCRX ligand/GPCRX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the GPCRX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-GPCRX antibodies in a subject, to purify GPCRX ligands, and in screening assays to identify molecules that inhibit the interaction of GPCRX with an GPCRX ligand.

An GPCRX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.,* Ausubel, *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An GPCRX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the GPCRX protein.

GPCRX Agonists and Antagonists

The invention also pertains to variants of the GPCRX proteins that function as either GPCRX agonists (*i.e.,* mimetics) or as GPCRX antagonists. Variants of the GPCRX protein can be generated by mutagenesis (*e.g.*, discrete point mutation or truncation of the GPCRX protein). An agonist of the GPCRX protein can retain substantially the same, or a subset of,

the biological activities of the naturally occurring form of the GPCRX protein. An antagonist of the GPCRX protein can inhibit one or more of the activities of the naturally occurring form of the GPCRX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the GPCRX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the GPCRX proteins.

Variants of the GPCRX proteins that function as either GPCRX agonists (*i.e.*, mimetics) or as GPCRX antagonists can be identified by screening combinatorial libraries of mutants (*e.g.*, truncation mutants) of the GPCRX proteins for GPCRX protein agonist or antagonist activity. In one embodiment, a variegated library of GPCRX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of GPCRX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential GPCRX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of GPCRX sequences therein. There are a variety of methods which can be used to produce libraries of potential GPCRX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential GPCRX sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. *See, e.g.*, Narang, 1983. *Tetrahedron* 39: 3; Itakura, *et al.*, 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, *et al.*, 1984. *Science* 198: 1056; Ike, *et al.*, 1983. *Nucl. Acids Res.* 11: 477.

Polypeptide Libraries

In addition, libraries of fragments of the GPCRX protein coding sequences can be used to generate a variegated population of GPCRX fragments for screening and subsequent selection of variants of an GPCRX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an GPCRX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded

DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S_1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the GPCRX proteins.

Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of GPCRX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify GPCRX variants. *See, e.g.,* Arkin and Yourvan, 1992. *Proc. Natl. Acad. Sci. USA* 89: 7811-7815; Delgrave, *et al.*, 1993. *Protein Engineering* 6:327-331.

Anti-GPCRX Antibodies

Also included in the invention are antibodies to GPCRX proteins, or fragments of GPCRX proteins. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated GPCRX-related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for

polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of GPCRX-related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human GPCRX-related protein sequence will indicate which regions of a GPCRX-related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow and Lane, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native

protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to

elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, *J. Immunol.*, 133:3001 (1984); Brodeur et al., MONOCLONAL ANTIBODY PRODUCTION TECHNIQUES AND APPLICATIONS, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, *Anal. Biochem.*, 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, *Nature* 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al.,

Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)).

Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (*Bio/Technology* 10, 779-783 (1992)); Lonberg et al. (*Nature* 368 856-859 (1994)); Morrison (*Nature* 368, 812-13 (1994)); Fishwild

et al,(*Nature Biotechnology* 14, 845-51 (1996)); Neuberger (*Nature Biotechnology* 14, 826 (1996)); and Lonberg and Huszar (*Intern. Rev. Immunol.* 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into

another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, *Nature*, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion

preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., *Methods in Enzymology*, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., *Science* 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to

cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., *J. Immunol.* 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., *J. Immunol.* 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells

(U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of an GPCRX protein is facilitated by generation of hybridomas that bind to the fragment of an GPCRX protein possessing such a domain. Thus, antibodies that are specific for a desired domain within an GPCRX protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-GPCRX antibodies may be used in methods known within the art relating to the localization and/or quantitation of an GPCRX protein (*e.g.*, for use in measuring levels of the GPCRX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for GPCRX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds (hereinafter "Therapeutics").

An anti-GPCRX antibody (*e.g.*, monoclonal antibody) can be used to isolate an GPCRX polypeptide by standard techniques, such as affinity chromatography or

immunoprecipitation. An anti-GPCRX antibody can facilitate the purification of natural GPCRX polypeptide from cells and of recombinantly-produced GPCRX polypeptide expressed in host cells. Moreover, an anti-GPCRX antibody can be used to detect GPCRX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the GPCRX protein. Anti-GPCRX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

GPCRX Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an GPCRX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to

include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell).

The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, GPCRX proteins, mutant forms of GPCRX proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of GPCRX proteins in prokaryotic or eukaryotic cells. For example, GPCRX proteins can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors

typically serve three purposes: (*i*) to increase expression of recombinant protein; (*ii*) to increase the solubility of the recombinant protein; and (*iii*) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. *See, e.g.,* Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (*see, e.g.,* Wada, *et al.*, 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the GPCRX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerivisae* include pYepSec1 (Baldari, *et al.*, 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz *et al.*, 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, GPCRX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith, *et al.*, 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors

include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, *et al.*, 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, and simian virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see, *e.g.*, Chapters 16 and 17 of Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, *et al.*, 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Banerji, *et al.*, 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, *et al.*, 1985. *Science* 230: 912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the α-fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to GPCRX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene

expression using antisense genes *see, e.g.,* Weintraub, *et al.*, "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, GPCRX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding GPCRX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) GPCRX protein. Accordingly, the invention further provides

methods for producing GPCRX protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding GPCRX protein has been introduced) in a suitable medium such that GPCRX protein is produced. In another embodiment, the method further comprises isolating GPCRX protein from the medium or the host cell.

Transgenic GPCRX Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which GPCRX protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous GPCRX sequences have been introduced into their genome or homologous recombinant animals in which endogenous GPCRX sequences have been altered. Such animals are useful for studying the function and/or activity of GPCRX protein and for identifying and/or evaluating modulators of GPCRX protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous GPCRX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing GPCRX-encoding nucleic acid into the male pronuclei of a fertilized oocyte (*e.g.*, by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The human GPCRX cDNA sequences of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human GPCRX gene, such as a mouse GPCRX gene, can be isolated based on hybridization to the human GPCRX cDNA (described further *supra*) and used as a transgene. Intronic sequences and polyadenylation signals can also be

included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the GPCRX transgene to direct expression of GPCRX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the GPCRX transgene in its genome and/or expression of GPCRX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding GPCRX protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of an GPCRX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the GPCRX gene. The GPCRX gene can be a human gene (*e.g.*, the cDNA of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27), but more preferably, is a non-human homologue of a human GPCRX gene. For example, a mouse homologue of human GPCRX gene of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 can be used to construct a homologous recombination vector suitable for altering an endogenous GPCRX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous GPCRX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous GPCRX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous GPCRX protein). In the homologous recombination vector, the altered portion of the GPCRX gene is flanked at its 5'- and 3'-termini by additional nucleic acid of the GPCRX gene to allow for homologous recombination to occur between the exogenous GPCRX gene carried by the vector and an endogenous GPCRX gene in an embryonic stem cell. The additional flanking GPCRX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5'- and 3'-termini) are included in the vector. *See, e.g.,* Thomas, *et*

al., 1987. *Cell* 51: 503 for a description of homologous recombination vectors. The vector is ten introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced GPCRX gene has homologously-recombined with the endogenous GPCRX gene are selected. *See, e.g.*, Li, *et al.*, 1992. *Cell* 69: 915.

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. *See, e.g.*, Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, 1991. *Curr. Opin. Biotechnol.* 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, *See, e.g.*, Lakso, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae*. *See,* O'Gorman, *et al.*, 1991. *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, *et al.*, 1997. *Nature* 385: 810-813. In brief, a cell (*e.g.*, a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell (*e.g.*, the somatic cell) is isolated.

Pharmaceutical Compositions

The GPCRX nucleic acid molecules, GPCRX proteins, and anti-GPCRX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (*i.e.*, topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration,

suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, an GPCRX protein or anti-GPCRX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient

such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent

on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.,* U.S. Patent No. 5,328,470) or by stereotactic injection (*see, e.g.*, Chen, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

Screening and Detection Methods

The isolated nucleic acid molecules of the invention can be used to express GPCRX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect GPCRX mRNA (*e.g.*, in a biological sample) or a genetic lesion in an GPCRX gene, and to modulate GPCRX activity, as described further, below. In addition, the GPCRX proteins can be used to screen drugs or compounds that modulate the GPCRX protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of GPCRX protein or production of GPCRX protein forms that have decreased or aberrant activity compared to GPCRX wild-type protein (*e.g.*; diabetes (regulates insulin release); obesity (binds and transport lipids); metabolic disturbances associated with obesity, the metabolic syndrome X as well as anorexia and wasting disorders associated with chronic diseases and various cancers, and infectious disease(possesses anti-microbial activity) and the various dyslipidemias. In addition, the anti-GPCRX antibodies of the invention can be used to detect and isolate GPCRX proteins and modulate GPCRX activity. In yet a further aspect, the invention can be used in methods to influence appetite, absorption of nutrients and the disposition of metabolic substrates in both a positive and negative fashion.

The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, *supra*.

Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to GPCRX proteins or have a stimulatory or inhibitory effect on, *e.g.*, GPCRX protein expression or GPCRX protein activity. The invention also includes compounds identified in the screening assays described herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of an GPCRX protein or polypeptide or biologically-active portion thereof. The test compounds of the invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds. *See, e.g.,* Lam, 1997. *Anticancer Drug Design* 12: 145.

A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, *e.g.*, nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and/or biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art and can be screened with any of the assays of the invention.

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt, *et al.*, 1993. *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb, *et al.*, 1994. *Proc. Natl. Acad. Sci. U.S.A.* 91: 11422; Zuckermann, *et al.*, 1994. *J. Med. Chem.* 37: 2678; Cho, *et al.*, 1993. *Science* 261: 1303; Carrell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2059; Carell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2061; and Gallop, *et al.*, 1994. *J. Med. Chem.* 37: 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992. *Biotechniques* 13: 412-421), or on beads (Lam, 1991. *Nature* 354: 82-84), on chips (Fodor, 1993. *Nature* 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 1865-1869) or on phage (Scott and Smith, 1990. *Science* 249: 386-390; Devlin, 1990. *Science*

249: 404-406; Cwirla, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6378-6382; Felici, 1991. *J. Mol. Biol.* 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to an GPCRX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the GPCRX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the GPCRX protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically-labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds GPCRX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the test compound to preferentially bind to GPCRX protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of GPCRX protein, or a biologically-active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the GPCRX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCRX or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule. As used herein, a "target molecule" is a molecule with which an GPCRX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses an GPCRX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. An GPCRX target molecule can be a non-GPCRX

molecule or an GPCRX protein or polypeptide of the invention. In one embodiment, an GPCRX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.* a signal generated by binding of a compound to a membrane-bound GPCRX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with GPCRX.

Determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the GPCRX protein to bind to or interact with an GPCRX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+}, diacylglycerol, IP_3, etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising an GPCRX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting an GPCRX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to bind to the GPCRX protein or biologically-active portion thereof. Binding of the test compound to the GPCRX protein can be determined either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the GPCRX protein or biologically-active portion thereof with a known compound which binds GPCRX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the test compound to preferentially bind to GPCRX or biologically-active portion thereof as compared to the known compound.

In still another embodiment, an assay is a cell-free assay comprising contacting GPCRX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the GPCRX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of GPCRX can be accomplished, for example, by determining the ability of the GPCRX protein to bind to an GPCRX target molecule by one of

the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of GPCRX protein can be accomplished by determining the ability of the GPCRX protein further modulate an GPCRX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as described, *supra*.

In yet another embodiment, the cell-free assay comprises contacting the GPCRX protein or biologically-active portion thereof with a known compound which binds GPCRX protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an GPCRX protein, wherein determining the ability of the test compound to interact with an GPCRX protein comprises determining the ability of the GPCRX protein to preferentially bind to or modulate the activity of an GPCRX target molecule.

The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of GPCRX protein. In the case of cell-free assays comprising the membrane-bound form of GPCRX protein, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of GPCRX protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the invention, it may be desirable to immobilize either GPCRX protein or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to GPCRX protein, or interaction of GPCRX protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-GPCRX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or GPCRX protein, and the mixture

is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, *supra*. Alternatively, the complexes can be dissociated from the matrix, and the level of GPCRX protein binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the GPCRX protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated GPCRX protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with GPCRX protein or target molecules, but which do not interfere with binding of the GPCRX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or GPCRX protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the GPCRX protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the GPCRX protein or target molecule.

In another embodiment, modulators of GPCRX protein expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of GPCRX mRNA or protein in the cell is determined. The level of expression of GPCRX mRNA or protein in the presence of the candidate compound is compared to the level of expression of GPCRX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of GPCRX mRNA or protein expression based upon this comparison. For example, when expression of GPCRX mRNA or protein is greater (*i.e.*, statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of GPCRX mRNA or protein expression. Alternatively, when expression of GPCRX mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of GPCRX mRNA or protein expression. The level of GPCRX mRNA or protein expression in the cells can be determined by methods described herein for detecting GPCRX mRNA or protein.

In yet another aspect of the invention, the GPCRX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Bartel, *et al.*, 1993. *Biotechniques* 14: 920-924; Iwabuchi, *et al.*, 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with GPCRX ("GPCRX-binding proteins" or "GPCRX-bp") and modulate GPCRX activity. Such GPCRX-binding proteins are also likely to be involved in the propagation of signals by the GPCRX proteins as, for example, upstream or downstream elements of the GPCRX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for GPCRX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an GPCRX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with GPCRX.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, these sequences can be used to: (*i*) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (*ii*) identify an individual from a minute biological sample (tissue typing); and (*iii*) aid in forensic identification of a biological sample. Some of these applications are described in the subsections, below.

Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the GPCRX sequences, SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or fragments or derivatives thereof, can be used to map the location of the GPCRX genes, respectively, on a chromosome. The mapping of the GPCRX sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, GPCRX genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the GPCRX sequences. Computer analysis of the GPCRX, sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the GPCRX sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. *See, e.g.,* D'Eustachio, *et al.*, 1983. *Science* 220: 919-924. Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the GPCRX sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in

metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases, will suffice to get good results at a reasonable amount of time. For a review of this technique, *see,* Verma, *et al.*, HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, *e.g.,* in McKusick, MENDELIAN INHERITANCE IN MAN, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, *e.g.,* Egeland, *et al.*, 1987. *Nature*, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the GPCRX gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

Tissue Typing

The GPCRX sequences of the invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with

one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Patent No. 5,272,057).

Furthermore, the sequences of the invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the GPCRX sequences described herein can be used to prepare two PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the invention can be used to obtain such identification sequences from individuals and from tissue. The GPCRX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the invention relates to diagnostic assays for determining GPCRX protein and/or nucleic acid expression as well as GPCRX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant GPCRX

expression or activity. The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with GPCRX protein, nucleic acid expression or activity. For example, mutations in an GPCRX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with GPCRX protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining GPCRX protein, nucleic acid expression or activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of GPCRX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

An exemplary method for detecting the presence or absence of GPCRX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting GPCRX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes GPCRX protein such that the presence of GPCRX is detected in the biological sample. An agent for detecting GPCRX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to GPCRX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length GPCRX nucleic acid, such as the nucleic acid of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to GPCRX mRNA or

genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting GPCRX protein is an antibody capable of binding to GPCRX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect GPCRX mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of GPCRX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of GPCRX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of GPCRX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of GPCRX protein include introducing into a subject a labeled anti-GPCRX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting GPCRX protein, mRNA, or genomic DNA, such that the presence of GPCRX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of GPCRX protein, mRNA or genomic DNA in the control sample with the presence of GPCRX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of GPCRX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting GPCRX protein or mRNA in a biological sample; means for determining the amount of GPCRX in the sample; and means for comparing the amount of GPCRX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect GPCRX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant GPCRX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with GPCRX protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant GPCRX expression or activity in which a test sample is obtained from a subject and GPCRX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of GPCRX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant GPCRX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant GPCRX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder. Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant GPCRX expression or activity in which a test sample is obtained and GPCRX protein or nucleic acid is detected (*e.g.*, wherein the presence of GPCRX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant GPCRX expression or activity).

The methods of the invention can also be used to detect genetic lesions in an GPCRX gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding an GPCRX-protein, or the misexpression of the GPCRX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of: (*i*) a deletion of one or more nucleotides from an GPCRX gene; (*ii*) an addition of one or more nucleotides to an GPCRX gene; (*iii*) a substitution of one or more nucleotides of an GPCRX gene, (*iv*) a chromosomal rearrangement of an GPCRX gene; (*v*) an alteration in the level of a messenger RNA transcript of an GPCRX gene, (*vi*) aberrant modification of an GPCRX gene, such as of the methylation pattern of the genomic DNA, (*vii*) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of an GPCRX gene, (*viii*) a non-wild-type level of an GPCRX protein, (*ix*) allelic loss of an GPCRX gene, and (*x*) inappropriate post-translational modification of an GPCRX protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in an GPCRX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran, *et al.*, 1988. *Science* 241: 1077-1080; and Nakazawa, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 360-364), the latter of which can be particularly useful for detecting point mutations in the GPCRX-gene (*see,* Abravaya, *et al.*, 1995. *Nucl. Acids Res.* 23: 675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to an GPCRX gene under conditions such that hybridization and amplification of the GPCRX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (*see,* Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), transcriptional amplification system (*see,* Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177); Qβ Replicase (*see,* Lizardi, *et al,* 1988. *BioTechnology* 6: 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in an GPCRX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.,* U.S. Patent No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in GPCRX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high-density arrays containing hundreds or thousands of oligonucleotides probes. *See, e.g.,* Cronin, *et al.*, 1996. *Human Mutation* 7: 244-255; Kozal, *et al.*, 1996. *Nat. Med.* 2: 753-759. For example, genetic mutations in GPCRX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, *et al.*, *supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the GPCRX gene and detect mutations by comparing the sequence of the sample GPCRX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert, 1977. *Proc. Natl. Acad. Sci. USA* 74: 560 or Sanger, 1977. *Proc. Natl. Acad. Sci. USA* 74: 5463. It is also contemplated that any of a variety of automated sequencing procedures

can be utilized when performing the diagnostic assays (*see, e.g.,* Naeve, *et al.*, 1995. *Biotechniques* 19: 448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen, *et al.*, 1996. *Adv. Chromatography* 36: 127-162; and Griffin, *et al.*, 1993. *Appl. Biochem. Biotechnol.* 38: 147-159).

Other methods for detecting mutations in the GPCRX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes. *See, e.g.,* Myers, *et al.*, 1985. *Science* 230: 1242. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type GPCRX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S_1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g.,* Cotton, *et al.*, 1988. *Proc. Natl. Acad. Sci. USA* 85: 4397; Saleeba, *et al.*, 1992. *Methods Enzymol.* 217: 286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in GPCRX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches. *See, e.g.,* Hsu, *et al.*, 1994. *Carcinogenesis* 15: 1657-1662. According to an exemplary embodiment, a probe based on an GPCRX sequence, *e.g.*, a wild-type GPCRX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. *See, e.g.,* U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in GPCRX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids. *See, e.g.,* Orita, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA*: 86: 2766; Cotton,

1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79. Single-stranded DNA fragments of sample and control GPCRX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. *See, e.g.,* Keen, *et al.*, 1991. *Trends Genet.* 7: 5.

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). *See, e.g.,* Myers, *et al.,* 1985. *Nature* 313: 495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. *See, e.g.,* Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. *See, e.g.,* Saiki, *et al.*, 1986. *Nature* 324: 163; Saiki, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; *see, e.g.,* Gibbs, *et al.*, 1989. *Nucl. Acids Res.* 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (*see, e.g.,* Prossner, 1993. *Tibtech.* 11: 238). In addition it may be desirable to introduce a novel

restriction site in the region of the mutation to create cleavage-based detection. *See, e.g.,* Gasparini, *et al.,* 1992. *Mol. Cell Probes* 6: 1. It is anticipated that in certain embodiments amplification may also be performed using *Taq* ligase for amplification. *See, e.g.,* Barany, 1991. *Proc. Natl. Acad. Sci. USA* 88: 189. In such cases, ligation will occur only if there is a perfect match at the 3'-terminus of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.,* in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving an GPCRX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which GPCRX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on GPCRX activity (*e.g.,* GPCRX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.) In conjunction with such treatment, the pharmacogenomics (*i.e.,* the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.,* drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of GPCRX protein, expression of GPCRX nucleic acid, or mutation content of GPCRX genes in an

individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996. *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of GPCRX protein, expression of GPCRX nucleic acid, or mutation content of GPCRX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness

phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with an GPCRX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of GPCRX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase GPCRX gene expression, protein levels, or upregulate GPCRX activity, can be monitored in clinical trails of subjects exhibiting decreased GPCRX gene expression, protein levels, or downregulated GPCRX activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease GPCRX gene expression, protein levels, or downregulate GPCRX activity, can be monitored in clinical trails of subjects exhibiting increased GPCRX gene expression, protein levels, or upregulated GPCRX activity. In such clinical trials, the expression or activity of GPCRX and, preferably, other genes that have been implicated in, for example, a cellular proliferation or immune disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

By way of example, and not of limitation, genes, including GPCRX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates GPCRX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of GPCRX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of GPCRX or other genes. In this manner, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the

screening assays described herein) comprising the steps of (*i*) obtaining a pre-administration sample from a subject prior to administration of the agent; (*ii*) detecting the level of expression of an GPCRX protein, mRNA, or genomic DNA in the preadministration sample; (*iii*) obtaining one or more post-administration samples from the subject; (*iv*) detecting the level of expression or activity of the GPCRX protein, mRNA, or genomic DNA in the post-administration samples; (*v*) comparing the level of expression or activity of the GPCRX protein, mRNA, or genomic DNA in the pre-administration sample with the GPCRX protein, mRNA, or genomic DNA in the post administration sample or samples; and (*vi*) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of GPCRX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of GPCRX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant GPCRX expression or activity. The disorders include cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia, prostate cancer, neoplasm; adenocarcinoma, lymphoma, uterus cancer, fertility, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial asthma, Crohn's disease; multiple sclerosis, treatment of Albright Hereditary Ostoeodystrophy, and other diseases, disorders and conditions of the like.

These methods of treatment will be discussed more fully, below.

Disease and Disorders

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to: (*i*) an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; (*ii*) antibodies to an aforementioned peptide; (*iii*)

nucleic acids encoding an aforementioned peptide; (*iv*) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to an aforementioned peptide) that are utilized to "knockout" endoggenous function of an aforementioned peptide by homologous recombination (*see, e.g.,* Capecchi, 1989. *Science* 244: 1288-1292); or (*v*) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between an aforementioned peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.,* from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of an aforementioned peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.,* by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.,* Northern assays, dot blots, *in situ* hybridization, and the like).

Prophylactic Methods

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant GPCRX expression or activity, by administering to the subject an agent that modulates GPCRX expression or at least one GPCRX activity. Subjects at risk for a disease that is caused or contributed to by aberrant GPCRX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the GPCRX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of GPCRX aberrancy, for example, an GPCRX agonist or GPCRX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays

described herein. The prophylactic methods of the invention are further discussed in the following subsections.

Therapeutic Methods

Another aspect of the invention pertains to methods of modulating GPCRX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of GPCRX protein activity associated with the cell. An agent that modulates GPCRX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of an GPCRX protein, a peptide, an GPCRX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more GPCRX protein activity. Examples of such stimulatory agents include active GPCRX protein and a nucleic acid molecule encoding GPCRX that has been introduced into the cell. In another embodiment, the agent inhibits one or more GPCRX protein activity. Examples of such inhibitory agents include antisense GPCRX nucleic acid molecules and anti-GPCRX antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of an GPCRX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, up-regulates or down-regulates) GPCRX expression or activity. In another embodiment, the method involves administering an GPCRX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant GPCRX expression or activity.

Stimulation of GPCRX activity is desirable in situations in which GPCRX is abnormally downregulated and/or in which increased GPCRX activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and/or differentiation (*e.g.*, cancer or immune associated disorders). Another example of such a situation is where the subject has a gestational disease (*e.g.*, preclampsia).

Determination of the Biological Effect of the Therapeutic

In various embodiments of the invention, suitable *in vitro* or *in vivo* assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Prophylactic and Therapeutic Uses of the Compositions of the Invention

The GPCRX nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including, but not limited to: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.

As an example, a cDNA encoding the GPCRX protein of the invention may be useful in gene therapy, and the protein may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will have efficacy for treatment of patients suffering from: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias.

Both the novel nucleic acid encoding the GPCRX protein, and the GPCRX protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (*i.e.,* some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods.

Examples

Example 1. Identification of GPCRX clones

The novel nucleic acid sequences of GPCR1 through GPCR5 were identified on chromosome 11 by TblastN using CuraGen Corporation's sequence files for Olfactory Receptor homolog, run against the Genomic Daily Files made available by GenBank. The 165 kbp human genomic clone from CuraGen acc:AP0010804HTG derived from Homo sapiens chromosome 11, clone RP11-164A10 map 11q, was analyzed by GenScan and Grail software to identify exons and putative coding sequences. These clones were also analyzed by TblastN, BlastX and other programs to identify genomic regions translating to proteins with similarity to the original protein or protein family of interest.

All novel GPCRX target sequences identified in the present invention were subjected to the exon linking process to confirm the sequence. PCR primers were designed by starting at the most upstream sequence available, for the forward primer, and at the most downstream sequence available for the reverse primer. In each case, the sequence was examined, walking inward from the respective termini toward the coding sequence, until a suitable sequence that is either unique or highly selective was encountered, or, in the case of the reverse primer, until the stop codon was reached. Such primers were designed based on in silico predictions for the full length cDNA, part (one or more exons) of the DNA or protein sequence of the target sequence, or by translated homology of the predicted exons to closely related human sequences from other species. These primers were then employed in PCR amplification based on the following pool of human cDNAs: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain -whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea, uterus. Usually the resulting amplicons were gel purified, cloned and sequenced to high redundancy. The resulting sequences from all clones were assembled with themselves, with other fragments in CuraGen Corporation's database and with public ESTs. Fragments and ESTs were included as components for an assembly when the extent of their identity with another component of the assembly was at least 95% over 50 bp. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. These procedures provide the sequence reported herein.

Example 2. Quantitative expression analysis of clones in various cells and tissues

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR; TAQMAN®). RTQ PCR was performed on a Perkin-Elmer Biosystems ABI PRISM® 7700 Sequence Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing cells and cell lines from normal and cancer sources), Panel 2 (containing samples derived from tissues, in particular from surgical samples, from normal and cancer sources), Panel 3 (containing samples derived from a wide variety of cancer sources), Panel 4 (containing cells and cell lines from normal cells and cells related to inflammatory conditions) and Panel CNSD.01 (containing samples from normal and diseased brains).

First, the RNA samples were normalized to constitutively expressed genes such as β-actin and GAPDH. RNA (~50 ng total or ~1 ng polyA+) was converted to cDNA using the TAQMAN® Reverse Transcription Reagents Kit (PE Biosystems, Foster City, CA; Catalog No. N808-0234) and random hexamers according to the manufacturer's protocol. Reactions were performed in 20 ul and incubated for 30 min. at 48°C. cDNA (5 ul) was then transferred to a separate plate for the TAQMAN® reaction using β-actin and GAPDH TAQMAN® Assay Reagents (PE Biosystems; Catalog Nos. 4310881E and 4310884E, respectively) and TAQMAN® universal PCR Master Mix (PE Biosystems; Catalog No. 4304447) according to the manufacturer's protocol. Reactions were performed in 25 ul using the following parameters: 2 min. at 50°C; 10 min. at 95°C; 15 sec. at 95°C/1 min. at 60°C (40 cycles). Results were recorded as CT values (cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying by 100. The average CT values obtained for ß-actin and GAPDH were used to normalize RNA samples. The RNA sample generating the highest CT value required no further diluting, while all other samples were diluted relative to this sample according to their β-actin /GAPDH average CT values.

Normalized RNA (5 ul) was converted to cDNA and analyzed via TAQMAN® using One Step RT-PCR Master Mix Reagents (PE Biosystems; Catalog No. 4309169) and gene-specific primers according to the manufacturer's instructions. Probes and primers were designed for each assay according to Perkin Elmer Biosystem's *Primer Express* Software

package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer concentration = 250 nM, primer melting temperature (T_m) range = 58°-60° C, primer optimal Tm = 59° C, maximum primer difference = 2° C, probe does not have 5' G, probe T_m must be 10° C greater than primer T_m, amplicon size 75 bp to 100 bp. The probes and primers selected (see below) were synthesized by Synthegen (Houston, TX, USA). Probes were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900 nM each, and probe, 200nM.

PCR conditions: Normalized RNA from each tissue and each cell line was spotted in each well of a 96 well PCR plate (Perkin Elmer Biosystems). PCR cocktails including two probes (a probe specific for the target clone and another gene-specific probe multiplexed with the target probe) were set up using 1X TaqMan™ PCR Master Mix for the PE Biosystems 7700, with 5 mM MgCl2, dNTPs (dA, G, C, U at 1:1:1:2 ratios), 0.25 U/ml AmpliTaq Gold™ (PE Biosystems), and 0.4 U/μl RNase inhibitor, and 0.25 U/μl reverse transcriptase. Reverse transcription was performed at 48° C for 30 minutes followed by amplification/PCR cycles as follows: 95° C 10 min, then 40 cycles of 95° C for 15 seconds, 60° C for 1 minute.

In the results for Panel 1, the following abbreviations are used:

ca. = carcinoma,
* = established from metastasis,
met = metastasis,
s cell var= small cell variant,
non-s = non-sm =non-small,
squam = squamous,
pl. eff = pl effusion = pleural effusion,
glio = glioma,
astro = astrocytoma, and
neuro = neuroblastoma.

Panel 2

The plates for Panel 2 generally include 2 control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by surgeons working in close cooperation with the National Cancer Institute's Cooperative Human Tissue Network (CHTN) or the National Disease Research Initiative (NDRI). The tissues are derived from human malignancies and in cases where indicated many malignant tissues have "matched margins"

obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologists at NDRI or CHTN). This analysis provides a gross histopathological assessment of tumor differentiation grade. Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (i.e. immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly people or sudden death victims (accidents, etc.). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

Panel 3D

The plates of Panel 3D are comprised of 94 cDNA samples and two control samples. Specifically, 92 of these samples are derived from cultured human cancer cell lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines are generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas, ovarian/uterine/cervical, gastric, colon, lung and CNS cancer cell lines. In addition, there are two independent samples of cerebellum. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. The cell lines in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be

indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

Panel 4

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4r) or cDNA (Panel 4d) isolated from various human cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene ,La Jolla, CA) and thymus and kidney (Clontech) were employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA). Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and/or 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-5 ng/ml, TNF alpha at approximately 5-10 ng/ml, IFN gamma at approximately 20-50 ng/ml, IL-4 at approximately 5-10 ng/ml, IL-9 at approximately 5-10 ng/ml, IL-13 at approximately 5-10 ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco/Life Technologies, Rockville, MD), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20 ng/ml PMA and 1-2 μg/ml ionomycin, IL-12 at 5-10 ng/ml, IFN gamma at 20-50 ng/ml and IL-18 at 5-10 ng/ml for 6 hours. In some cases, mononuclear cells were cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5 μg/ml. Samples

were taken at 24, 48 and 72 hours for RNA preparation. MLR (mixed lymphocyte reaction) samples were obtained by taking blood from two donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately $2x10^6$ cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol (5.5×10^{-5} M) (Gibco), and 10 mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +ve VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum (FCS) (Hyclone, Logan, UT), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco), 50 ng/ml GMCSF and 5 ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10 mM Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50 ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with lipopolysaccharide (LPS) at 100 ng/ml. Dendritic cells were also stimulated with anti-CD40 monoclonal antibody (Pharmingen) at 10 μg/ml for 6 and 12-14 hours.

CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56, CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. Then CD45RO beads were used to isolate the CD45RO CD4 lymphocytes with the remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and plated at 10^6 cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5 μg/ml anti-CD28 (Pharmingen) and 3 ug/ml anti-CD3 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested the cells and expanded them in

DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resupended at 10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco). To activate the cells, we used PWM at 5 μg/ml or anti-CD40 (Pharmingen) at approximately 10 μg/ml and IL-4 at 5-10 ng/ml. Cells were harvested for RNA preparation at 24,48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon plates were coated overnight with 10 μg/ml anti-CD28 (Pharmingen) and 2 μg/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at 10^5-10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), 10 mM Hepes (Gibco) and IL-2 (4 ng/ml). IL-12 (5 ng/ml) and anti-IL4 (1 μg/ml) were used to direct to Th1, while IL-4 (5 ng/ml) and anti-IFN gamma (1 μg/ml) were used to direct to Th2 and IL-10 at 5 ng/ml was used to direct to Tr1. After 4-5 days, the activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100 μM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), 10 mM Hepes (Gibco) and IL-2 (1 ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1 μg/ml) to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours following the second and

third activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1 mM dbcAMP at 5 $x10^5$ cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to 5 $x10^5$ cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100 µM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), 10 mM Hepes (Gibco). RNA was either prepared from resting cells or cells activated with PMA at 10 ng/ml and ionomycin at 1 µg/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100 µM non essential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), mercaptoethanol 5.5 x 10^{-5} M (Gibco), and 10 mM Hepes (Gibco). CCD1106 cells were activated for 6 and 14 hours with approximately 5 ng/ml TNF alpha and 1 ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5 ng/ml IL-4, 5 ng/ml IL-9, 5 ng/ml IL-13 and 25 ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately 10^7 cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15 ml Falcon Tube. An equal volume of isopropanol was added and left at –20 degrees C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet was redissolved in 300 µl of RNAse-free water and 35 µl buffer (Promega) 5 µl DTT, 7 µl RNAsin and 8 µl DNAse were added. The tube was incubated at 37 degrees C for 30 minutes to remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3 M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNAse free water. RNA was stored at –80 degrees C.

Panel CNSD.01

The plates for Panel CNSD.01 include two control wells and 94 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor.

All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains two brains from each of the following diagnoses: Alzheimer's disease, Parkinson's disease, Huntington's disease, Progressive Supernuclear Palsy, Depression, and "Normal controls". Within each of these brains, the following regions are represented: cingulate gyrus, temporal pole, globus palladus, substantia nigra, Brodman Area 4 (primary motor strip), Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area 17 (occipital cortex). Not all brain regions are represented in all cases; e.g., Huntington's disease is characterized in part by neurodegeneration in the globus palladus, thus this region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's disease is characterized by degeneration of the substantia nigra making this region more difficult to obtain. Normal control brains were examined for neuropathology and found to be free of any pathology consistent with neurodegeneration.

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

In the labels employed to identify tissues in the CNS panel, the following abbreviations are used:

PSP = Progressive supranuclear palsy

Sub Nigra = Substantia nigra

Glob Palladus= Globus palladus

Temp Pole = Temporal pole

Cing Gyr = Cingulate gyrus

BA 4 = Brodman Area 4

A. GPCR1 (also known as AP001804_A or CG54326-01)

Expression of gene AP001804_A was assessed using the primer-probe sets Ag1634 and Ag2357 (identical sequences), described in Table 12. Results of the RTQ-PCR runs are shown in Tables 13 and 14.

Table 12. Probe Name Ag1634/Ag2357

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-TGAACTTTGTTCCAGAGGAGAA-3'	59	22	248	80
Probe	TET-5'-TCTCCTTTCTGGAATGCATTACTCAA-3'-TAMRA	64.3	26	275	81
Reverse	5'-GGTAGCCTTCTGCAATTACAAA-3'	58.5	22	319	82

Table 13. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5589 _ag1634_b2	Tissue Name	Relative Expression(%) 1.3dx4tm5589 _ag1634_b2
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	4.0	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	0.0	Lung (fetal)	0.0
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	6.5
Brain (thalamus)	6.1	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	4.7	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	9.3
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
CNS ca. (glio/astro) U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	0.0
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	4.9
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	0.0
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.0	Mammary gland	0.0
CNS ca. (glio) U251	0.0	Breast ca.* (pl. effusion) MCF-7	28.7
CNS ca. (glio) SF-295	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	37.1
Heart	0.0	Breast ca. BT-549	0.0

Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	3.7	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	4.0
Thymus	0.0	Ovarian ca. OVCAR-4	0.0
Spleen	**100.0**	Ovarian ca. OVCAR-5	0.0
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	0.0	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	5.2
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	4.7
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0
Colon ca. HT29	4.5	Prostate ca.* (bone met)PC-3	4.2
Colon ca. HCT-116	0.0	Testis	9.3
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	23.7
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table 14. Panel 4D

Tissue Name	Relative Expression(%) 4dx4tm5519t_ ag1634_a2	Tissue Name	Relative Expression(%) 4dx4tm5519t_ ag1634_a2
93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.0	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.0	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting day 4-6 in IL-2	0.0	93781_HUVEC (Endothelial)_IL-11	0.0
93571_Secondary Tr1_resting day 4-6 in IL-2	0.0	93583_Lung Microvascular Endothelial Cells_none	0.0
93568_primary Th1_anti-CD28/anti-CD3	0.0	93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93569_primary Th2_anti-CD28/anti-CD3	0.0	92662_Microvascular Dermal endothelium_none	0.0
93570_primary Tr1_anti-CD28/anti-CD3	0.0	92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml)	0.0

93565_primary Th1_resting dy 4-6 in IL-2	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0
93567_primary Tr1_resting dy 4-6 in IL-2	0.0
93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3	0.0
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3	0.0
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.0
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	0.0
93354_CD4_none	0.0
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	0.0
93103_LAK cells_resting	0.0
93788_LAK cells_IL-2	0.0
93787_LAK cells_IL-2+IL-12	0.0
93789_LAK cells_IL-2+IFN gamma	0.0
93790_LAK cells_IL-2+ IL-18	0.0
93104_LAK cells_PMA/ionomycin and IL-18	0.0
93578_NK Cells IL-2_resting	0.0
93109_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93110_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93111_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93112_Mononuclear Cells (PBMCs)_resting	0.0
93113_Mononuclear Cells (PBMCs)_PWM	0.0
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0
93249_Ramos (B cell)_none	0.0
and IL1b (1 ng/ml)	
93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	0.0
93347_Small Airway Epithelium_none	0.0
93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
92668_Coronery Artery SMC_resting	0.0
92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93107_astrocytes_resting	0.0
93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
92666_KU-812 (Basophil)_resting	0.2
92667_KU-812 (Basophil)_PMA/ionoycin	0.4
93579_CCD1106 (Keratinocytes)_none	0.0
93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.0
93791_Liver Cirrhosis	5.3
93792_Lupus Kidney	0.0
93577_NCI-H292	0.0
93358_NCI-H292_IL-4	0.0
93360_NCI-H292_IL-9	0.0
93359_NCI-H292_IL-13	0.0
93357_NCI-H292_IFN gamma	0.0
93777_HPAEC_-	0.0
93778_HPAEC_IL-1 beta/TNA alpha	0.5
93254_Normal Human Lung Fibroblast_none	0.2
93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.0
93257_Normal Human Lung Fibroblast_IL-4	0.4
93256_Normal Human Lung	0.7

		Fibroblast_IL-9	
93250_Ramos (B cell)_ionomycin	0.0	93255_Normal Human Lung Fibroblast_IL-13	0.5
93349_B lymphocytes_PWM	0.0	93258_Normal Human Lung Fibroblast_IFN gamma	0.0
93350_B lymphoytes_CD40L and IL-4	0.0	93106_Dermal Fibroblasts CCD1070_resting	0.0
92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.0	93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	0.0
93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	0.0
93356_Dendritic Cells_none	0.2	93772_dermal fibroblast_IFN gamma	0.0
93355_Dendritic Cells_LPS 100 ng/ml	0.0	93771_dermal fibroblast_IL-4	0.0
93775_Dendritic Cells_anti-CD40	4.2	93259_IBD Colitis 1**	**100.0**
93774_Monocytes_resting	0.0	93260_IBD Colitis 2	0.8
93776_Monocytes_LPS 50 ng/ml	0.0	93261_IBD Crohns	0.0
93581_Macrophages_resting	2.7	735010_Colon_normal	7.3
93582_Macrophages_LPS 100 ng/ml	0.0	735019_Lung_none	0.4
93098_HUVEC (Endothelial)_none	0.0	64028-1_Thymus_none	0.0
93099_HUVEC (Endothelial)_starved	0.0	64030-1_Kidney_none	0.0

Panel 1.3D Summary: Ag1634 Expression of GPCR1 gene AP001804_A is low/undetectable (CT values >35) in all cell lines and tissues except for spleen. Therefore, this gene may be used to distinguish spleen from other tissues. Ag2357 Expression was low/undetectable (CT values 40) in all tissues tested.

Panel 2D Summary: Ag2357 Expression was low/undetectable (CT values 40) in all tissues tested and thus the results not shown.

Panel 2.2 Summary: Ag1634 Expression of gene AP001804_A is low/undetectable (CT values >35) in all cell lines and tissues on this panel thus the results not shown.

Panel 4D Summary: Ag1634 Expression of the AP001804_A transcript is detected in colitis 1 and in dendritic cells treated with anti-CD40. The protein encoded for by this antigen may be important in the inflammatory process and particularly in the function of activated dendritic cells. Antagonistic antibodies or small molecule therapeutics that inhibit AP001804_A protein function may therefore reduce or inhibit inflammation in the bowel due to inflammatory bowl disease (IBD). Ag2357 Expression was low/undetectable (CT values 40) in all tissues tested and chemistry control did not work well (CT = 35).

B. GPCR2 (also known as AP001804_B or CG54335-01)

Expression of gene AP001804_B was assessed using the primer-probe sets Ag2355 and Ag1635 (identical sequences), described in Table 15. Results of the RTQ-PCR runs are shown in Tables 16, 17, 18, and 19.

Table 15. Probe Name Ag2355/Ag1635

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-TCATACAAGTGCCATGATGAAA-3'	59	22	474	83
Probe	FAM-5'-TGTCCTTTTGCAAATCCCACATTATCA-3'-TAMRA	68	27	497	84
Reverse	5'-AGGGGAAGAACATCACAGAAGT-3'	59.1	22	530	85

Table 16. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5627f_ag2355_b1	Tissue Name	Relative Expression(%) 1.3dx4tm5627f_ag2355_b1
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	0.0	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	0.0	Lung (fetal)	0.0
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	0.0	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	0.0	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	0.0
CNS ca. (glio/astro) U87-MG	8.1	Lung ca. (non-sm. cell) A549	0.0
CNS ca. (glio/astro) U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	0.0
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	7.1
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	7.7
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0

CNS ca. (glio) SNB-19	6.4	Mammary gland	0.0
CNS ca. (glio) U251	0.0	Breast ca.* (pl. effusion) MCF-7	**100.0**
CNS ca. (glio) SF-295	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	57.3
Heart	0.0	Breast ca. BT-549	0.0
Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	6.1	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.0	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	0.0
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	0.0	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	19.2
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	4.7
Colon ca. HCT-116	0.0	Testis	53.1
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	10.9
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table 17. Panel 2D

Tissue Name	Relative Expression(%) 2dx4tm4923f_ag2355_a2	Tissue Name	Relative Expression(%) 2dx4tm4923f_ag2355_a2
Normal Colon GENPAK 061003	4.3	Kidney NAT Clontech 8120608	0.0
83219 CC Well to Mod Diff (ODO3866)	0.7	Kidney Cancer Clontech 8120613	0.0
83220 CC NAT (ODO3866)	3.7	Kidney NAT Clontech 8120614	0.4
83221 CC Gr.2 rectosigmoid (ODO3868)	0.0	Kidney Cancer Clontech 9010320	0.4
83222 CC NAT (ODO3868)	0.0	Kidney NAT Clontech 9010321	0.0
83235 CC Mod Diff (ODO3920)	0.0	Normal Uterus GENPAK 061018	1.0
83236 CC NAT (ODO3920)	0.3	Uterus Cancer GENPAK	0.0

83237 CC Gr.2 ascend colon (ODO3921)	0.3
83238 CC NAT (ODO3921)	1.4
83241 CC from Partial Hepatectomy (ODO4309)	0.0
83242 Liver NAT (ODO4309)	0.0
87472 Colon mets to lung (OD04451-01)	0.0
87473 Lung NAT (OD04451-02)	0.8
Normal Prostate Clontech A+ 6546-1	0.8
84140 Prostate Cancer (OD04410)	0.0
84141 Prostate NAT (OD04410)	0.0
87073 Prostate Cancer (OD04720-01)	0.0
87074 Prostate NAT (OD04720-02)	0.5
Normal Lung GENPAK 061010	1.8
83239 Lung Met to Muscle (ODO4286)	0.8
83240 Muscle NAT (ODO4286)	0.0
84136 Lung Malignant Cancer (OD03126)	0.0
84137 Lung NAT (OD03126)	0.0
84871 Lung Cancer (OD04404)	0.0
84872 Lung NAT (OD04404)	0.0
84875 Lung Cancer (OD04565)	0.0
84876 Lung NAT (OD04565)	0.0
85950 Lung Cancer (OD04237-01)	0.3
85970 Lung NAT (OD04237-02)	0.0
83255 Ocular Mel Met to Liver (ODO4310)	0.0
83256 Liver NAT (ODO4310)	0.0
84139 Melanoma Mets to Lung (OD04321)	0.0
84138 Lung NAT (OD04321)	0.0
064011	
Normal Thyroid Clontech A+ 6570-1	0.0
Thyroid Cancer GENPAK 064010	0.0
Thyroid Cancer INVITROGEN A302152	0.0
Thyroid NAT INVITROGEN A302153	0.0
Normal Breast GENPAK 061019	0.0
84877 Breast Cancer (OD04566)	0.0
85975 Breast Cancer (OD04590-01)	0.0
85976 Breast Cancer Mets (OD04590-03)	0.0
87070 Breast Cancer Metastasis (OD04655-05)	0.0
GENPAK Breast Cancer 064006	0.0
Breast Cancer Res. Gen. 1024	0.0
Breast Cancer Clontech 9100266	0.0
Breast NAT Clontech 9100265	0.4
Breast Cancer INVITROGEN A209073	0.0
Breast NAT INVITROGEN A2090734	0.4
Normal Liver GENPAK 061009	0.0
Liver Cancer GENPAK 064003	5.0
Liver Cancer Research Genetics RNA 1025	0.7
Liver Cancer Research Genetics RNA 1026	0.0
Paired Liver Cancer Tissue Research Genetics RNA 6004-T	0.3
Paired Liver Tissue Research Genetics RNA 6004-N	0.7
Paired Liver Cancer Tissue Research Genetics RNA 6005-T	0.0
Paired Liver Tissue Research Genetics RNA 6005-N	0.0
Normal Bladder GENPAK 061001	1.8
Bladder Cancer Research Genetics RNA 1023	0.0
Bladder Cancer INVITROGEN	18.1

		A302173	
Normal Kidney GENPAK 061008	0.9	87071 Bladder Cancer (OD04718-01)	0.0
83786 Kidney Ca, Nuclear grade 2 (OD04338)	2.1	87072 Bladder Normal Adjacent (OD04718-03)	0.0
83787 Kidney NAT (OD04338)	0.0	Normal Ovary Res. Gen.	0.0
83788 Kidney Ca Nuclear grade 1/2 (OD04339)	0.8	Ovarian Cancer GENPAK 064008	0.0
83789 Kidney NAT (OD04339)	0.8	87492 Ovary Cancer (OD04768-07)	**100.0**
83790 Kidney Ca, Clear cell type (OD04340)	0.0	87493 Ovary NAT (OD04768-08)	0.0
83791 Kidney NAT (OD04340)	0.0	Normal Stomach GENPAK 061017	0.2
83792 Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer Clontech 9060358	0.0
83793 Kidney NAT (OD04348)	0.0	NAT Stomach Clontech 9060359	0.0
87474 Kidney Cancer (OD04622-01)	0.0	Gastric Cancer Clontech 9060395	0.0
87475 Kidney NAT (OD04622-03)	0.0	NAT Stomach Clontech 9060394	0.0
85973 Kidney Cancer (OD04450-01)	0.0	Gastric Cancer Clontech 9060397	0.0
85974 Kidney NAT (OD04450-03)	0.0	NAT Stomach Clontech 9060396	0.0
Kidney Cancer Clontech 8120607	0.0	Gastric Cancer GENPAK 064005	0.7

Table 18. Panel 3D

Tissue Name	Relative Expression(%) 3dx4tm5123f_ag2355_a2	Tissue Name	Relative Expression(%) 3dx4tm5123f_ag2355_a2
94905_Daoy_Medulloblastoma/Cerebellum_sscDNA	0.0	94954_Ca Ski_Cervical epidermoid carcinoma (metastasis)_sscDNA	0.0
94906_TE671_Medulloblastom/Cerebellum_sscDNA	13.4	94955_ES-2_Ovarian clear cell carcinoma_sscDNA	0.0
94907_D283 Med_Medulloblastoma/Cerebellum_sscDNA	0.0	94957_Ramos/6h stim_"; Stimulated with PMA/ionomycin 6h_sscDNA	0.0
94908_PFSK-1_Primitive Neuroectodermal/Cerebellum_sscDNA	0.0	94958_Ramos/14h stim_"; Stimulated with PMA/ionomycin 14h_sscDNA	0.0
94909_XF-498_CNS_sscDNA	0.0	94962_MEG-01_Chronic myelogenous leukemia (megokaryoblast)_sscDNA	0.0
94910_SNB-78_CNS/glioma_sscDNA	0.0	94963_Raji_Burkitt's lymphoma_sscDNA	0.0

94911_SF-268_CNS/glioblastoma_sscDNA	0.0
94912_T98G_Glioblastoma_sscDNA	0.0
96776_SK-N-SH_Neuroblastoma (metastasis)_sscDNA	0.0
94913_SF-295_CNS/glioblastoma_sscDNA	0.0
94914_Cerebellum_sscDNA	0.0
96777_Cerebellum_sscDNA	0.0
94916_NCI-H292_Mucoepidermoid lung carcinoma_sscDNA	0.0
94917_DMS-114_Small cell lung cancer_sscDNA	57.8
94918_DMS-79_Small cell lung cancer/neuroendocrine_sscDNA	0.0
94919_NCI-H146_Small cell lung cancer/neuroendocrine_sscDNA	4.9
94920_NCI-H526_Small cell lung cancer/neuroendocrine_sscDNA	4.1
94921_NCI-N417_Small cell lung cancer/neuroendocrine_sscDNA	0.0
94923_NCI-H82_Small cell lung cancer/neuroendocrine_sscDNA	0.0
94924_NCI-H157_Squamous cell lung cancer (metastasis)_sscDNA	28.9
94925_NCI-H1155_Large cell lung cancer/neuroendocrine_sscDNA	0.0
94926_NCI-H1299_Large cell lung cancer/neuroendocrine_sscDNA	**100.0**
94927_NCI-H727_Lung carcinoid_sscDNA	0.0
94928_NCI-UMC-11_Lung carcinoid_sscDNA	47.1
94929_LX-1_Small cell lung cancer_sscDNA	0.0
94930_Colo-205_Colon cancer_sscDNA	0.0
94964_Daudi_Burkitt's lymphoma_sscDNA	0.0
94965_U266_B-cell plasmacytoma/myeloma_sscDNA	72.8
94968_CA46_Burkitt's lymphoma_sscDNA	0.0
94970_RL_non-Hodgkin's B-cell lymphoma_sscDNA	0.0
94972_JM1_pre-B-cell lymphoma/leukemia_sscDNA	0.0
94973_Jurkat_T cell leukemia_sscDNA	0.0
94974_TF-1_Erythroleukemia_sscDNA	0.0
94975_HUT 78_T-cell lymphoma_sscDNA	3.6
94977_U937_Histiocytic lymphoma_sscDNA	0.0
94980_KU-812_Myelogenous leukemia_sscDNA	0.0
94981_769-P_Clear cell renal carcinoma_sscDNA	0.0
94983_Caki-2_Clear cell renal carcinoma_sscDNA	0.0
94984_SW 839_Clear cell renal carcinoma_sscDNA	0.0
94986_G401_Wilms' tumor_sscDNA	0.0
94987_Hs766T_Pancreatic carcinoma (LN metastasis)_sscDNA	0.0
94988_CAPAN-1_Pancreatic adenocarcinoma (liver metastasis)_sscDNA	0.0
94989_SU86.86_Pancreatic carcinoma (liver metastasis)_sscDNA	0.0
94990_BxPC-3_Pancreatic adenocarcinoma_sscDNA	0.0
94991_HPAC_Pancreatic adenocarcinoma_sscDNA	0.0
94992_MIA PaCa-2_Pancreatic carcinoma_sscDNA	0.0

94931_KM12_Colon cancer_sscDNA	0.0	94993_CFPAC-1_Pancreatic ductal adenocarcinoma_sscDNA	0.0
94932_KM20L2_Colon cancer_sscDNA	0.0	94994_PANC-1_Pancreatic epithelioid ductal carcinoma_sscDNA	0.0
94933_NCI-H716_Colon cancer_sscDNA	0.0	94996_T24_Bladder carcinma (transitional cell)_sscDNA	0.0
94935_SW-48_Colon adenocarcinoma_sscDNA	0.0	94997_5637_Bladder carcinoma_sscDNA	0.0
94936_SW1116_Colon adenocarcinoma_sscDNA	0.0	94998_HT-1197_Bladder carcinoma_sscDNA	0.0
94937_LS 174T_Colon adenocarcinoma_sscDNA	0.0	94999_UM-UC-3_Bladder carcinma (transitional cell)_sscDNA	0.0
94938_SW-948_Colon adenocarcinoma_sscDNA	0.0	95000_A204_Rhabdomyosarcoma_sscDNA	0.0
94939_SW-480_Colon adenocarcinoma_sscDNA	0.0	95001_HT-1080_Fibrosarcoma_sscDNA	0.0
94940_NCI-SNU-5_Gastric carcinoma_sscDNA	0.0	95002_MG-63_Osteosarcoma (bone)_sscDNA	3.5
94941_KATO III_Gastric carcinoma_sscDNA	0.0	95003_SK-LMS-1_Leiomyosarcoma (vulva)_sscDNA	0.0
94943_NCI-SNU-16_Gastric carcinoma_sscDNA	0.0	95004_SJRH30_Rhabdomyosarcoma (met to bone marrow)_sscDNA	0.0
94944_NCI-SNU-1_Gastric carcinoma_sscDNA	0.0	95005_A431_Epidermoid carcinoma_sscDNA	0.0
94946_RF-1_Gastric adenocarcinoma_sscDNA	0.0	95007_WM266-4_Melanoma_sscDNA	5.6
94947_RF-48_Gastric adenocarcinoma_sscDNA	0.0	95010_DU 145_Prostate carcinoma (brain metastasis)_sscDNA	0.0
96778_MKN-45_Gastric carcinoma_sscDNA	0.0	95012_MDA-MB-468_Breast adenocarcinoma_sscDNA	0.0
94949_NCI-N87_Gastric carcinoma_sscDNA	6.9	95013_SCC-4_Squamous cell carcinoma of tongue_sscDNA	0.0
94951_OVCAR-5_Ovarian carcinoma_sscDNA	0.0	95014_SCC-9_Squamous cell carcinoma of tongue_sscDNA	0.0
94952_RL95-2_Uterine carcinoma_sscDNA	0.0	95015_SCC-15_Squamous cell carcinoma of tongue_sscDNA	0.0
94953_HelaS3_Cervical adenocarcinoma_sscDNA	2.7	95017_CAL 27_Squamous cell carcinoma of tongue_sscDNA	0.0

Table 19. Panel 4D

Tissue Name	Relative Expression(%) 4dx4tm4927f_ag2355_a2	Tissue Name	Relative Expression(%) 4dx4tm4927f_ag2355_a2

93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.0	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.0	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting day 4-6 in IL-2	0.0	93781_HUVEC (Endothelial)_IL-11	0.0
93571_Secondary Tr1_resting day 4-6 in IL-2	0.0	93583_Lung Microvascular Endothelial Cells_none	0.0
93568_primary Th1_anti-CD28/anti-CD3	0.0	93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93569_primary Th2_anti-CD28/anti-CD3	0.0	92662_Microvascular Dermal endothelium_none	0.0
93570_primary Tr1_anti-CD28/anti-CD3	0.0	92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93565_primary Th1_resting dy 4-6 in IL-2	0.0	93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0	93347_Small Airway Epithelium_none	0.0
93567_primary Tr1_resting dy 4-6 in IL-2	0.0	93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3	0.0	92668_Coronery Artery SMC_resting	0.0
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3	0.0	92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml)	14.7
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0	93107_astrocytes_resting	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.0	93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	0.0	92666_KU-812 (Basophil)_resting	0.0
93354_CD4_none	0.0	92667_KU-812 (Basophil)_PMA/ionoycin	22.3
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	0.0	93579_CCD1106 (Keratinocytes)_none	0.0
93103_LAK cells_resting	0.0	93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.0
93788_LAK cells_IL-2	0.0	93791_Liver Cirrhosis	71.5
93787_LAK cells_IL-2+IL-12	0.0	93792_Lupus Kidney	0.0
93789_LAK cells_IL-2+IFN	0.0	93577_NCI-H292	0.0

gamma			
93790_LAK cells_IL-2+ IL-18	0.0	93358_NCI-H292_IL-4	0.0
93104_LAK cells_PMA/ionomycin and IL-18	0.0	93360_NCI-H292_IL-9	0.0
93578_NK Cells IL-2_resting	0.0	93359_NCI-H292_IL-13	0.0
93109_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93357_NCI-H292_IFN gamma	0.0
93110_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93777_HPAEC_-	0.0
93111_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93778_HPAEC_IL-1 beta/TNA alpha	0.0
93112_Mononuclear Cells (PBMCs)_resting	0.0	93254_Normal Human Lung Fibroblast_none	34.1
93113_Mononuclear Cells (PBMCs)_PWM	0.0	93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.0
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0	93257_Normal Human Lung Fibroblast_IL-4	0.0
93249_Ramos (B cell)_none	0.0	93256_Normal Human Lung Fibroblast_IL-9	0.0
93250_Ramos (B cell)_ionomycin	0.0	93255_Normal Human Lung Fibroblast_IL-13	0.0
93349_B lymphocytes_PWM	0.0	93258_Normal Human Lung Fibroblast_IFN gamma	13.0
93350_B lymphoytes_CD40L and IL-4	0.0	93106_Dermal Fibroblasts CCD1070_resting	0.0
92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.0	93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	0.0
93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	0.0
93356_Dendritic Cells_none	0.0	93772_dermal fibroblast_IFN gamma	0.0
93355_Dendritic Cells_LPS 100 ng/ml	0.0	93771_dermal fibroblast_IL-4	0.0
93775_Dendritic Cells_anti-CD40	45.9	93259_IBD Colitis 1**	10.0
93774_Monocytes_resting	0.0	93260_IBD Colitis 2	12.4
93776_Monocytes_LPS 50 ng/ml	0.0	93261_IBD Crohns	16.6
93581_Macrophages_resting	35.6	735010_Colon_normal	**100.0**
93582_Macrophages_LPS 100 ng/ml	0.0	735019_Lung_none	43.0
93098_HUVEC (Endothelial)_none	0.0	64028-1_Thymus_none	0.0
93099_HUVEC (Endothelial)_starved	0.0	64030-1_Kidney_none	0.0

Panel 1.3D Summary: Ag1635 The expression of GPCR2 gene AP001804_B is low/undetectable (CT values >35) in all the tissues on this panel. Ag2355 The expression of

the AP001804_B gene is low but significant in two breast cancer cell lines. Interestingly, the two positive breast cancer cell lines are estrogen receptor positive. Thus, expression of this gene may be indicative of estrogen receptor status on breast cancer cells and may have implications to breast cancer cell biology. In addition, therapeutic modulation of this gene may have utility in the treatment of breast cancer or other breast disease.

Panel 2 Summary: Ag2355 Expression of this gene is highest in a sample derived from an ovarian cancer. Samples in which there is also expression are many fold lower than the ovarian cancer. Thus, this gene may be useful for the diagnosis or therapeutic intervention for ovarian cancer.

Panel 2.2 Summary: Ag1635 Expression of gene AP001804_B on this panel is too low to be reliable (Ct values >35).

Panel 3D Summary: Ag2355 The expression of the AP001804_B gene in panel 3D appears to be associated with lung cancer cell lines. Furthermore, the cell line that expresses this gene in most abundance is neuroendocrine in origin. Neuroendocrine tumors are very unique and thus, the AP001804_B gene may represent a unique marker of this type of cancer. In addition, therapeutic modulation of this gene may be useful for the treatment of neuroendocrine tumors in the lung.

Panel 4D Summary: Ag1635 The AP001804_B transcript is expressed in normal colon but not in colons from patients with Crohn's disease or colitis. Protein therapeutics designed with the putative GPCR encoded for by this gene could be used to inhibit inflammation and tissue destruction due to IBD.

C. GPCR3 (also known as AP001804_C or CG54344-01)

Expression of gene AP001804_C was assessed using the primer-probe set Ag1639, described in Table 20. Results of the RTQ-PCR runs are shown in Tables 21, 22, and 23.

Table 20. Probe Name Ag1639

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-AGCATCTTCCACATTGATTCC-3'	59	21	670	86
Probe	TET-5'-CTTCAGCACCTGCAGCTCCCACATAA-3'-TAMRA	71.2	26	711	87
Reverse	5'-CCAAAGAACAGAGAAACTGCAA-3'	59.5	22	737	88

Table 21. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5590t_ag1639_a2	Tissue Name	Relative Expression(%) 1.3dx4tm5590t_ag1639_a2
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	17.2
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	0.0	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	0.0	Lung (fetal)	0.0
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	0.0	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	6.7	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	0.0
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
CNS ca. (glio/astro) U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	0.0
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	0.0
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	0.0
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.0	Mammary gland	0.0
CNS ca. (glio) U251	10.6	Breast ca.* (pl. effusion) MCF-7	15.2
CNS ca. (glio) SF-295	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	25.7
Heart	0.0	Breast ca. BT-549	0.0
Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.0	Ovarian ca. OVCAR-4	0.0
Spleen	**100.0**	Ovarian ca. OVCAR-5	0.0
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	0.0	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	6.6
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0

Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	6.1
Colon ca. HCT-116	0.0	Testis	0.0
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	37.4
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table 22. Panel 2.2

Tissue Name	Relative Expression(%) 2.2x4tm6361t_ ag1639_a2	Tissue Name	Relative Expression(%) 2.2x4tm6361t_ ag1639_a2
Normal Colon GENPAK 061003	17.6	83793 Kidney NAT (OD04348)	0.0
97759 Colon cancer (OD06064)	0.0	98938 Kidney malignant cancer (OD06204B)	0.0
97760 Colon cancer NAT (OD06064)	0.0	98939 Kidney normal adjacent tissue (OD06204E)	0.0
97778 Colon cancer (OD06159)	0.0	85973 Kidney Cancer (OD04450-01)	0.0
97779 Colon cancer NAT (OD06159)	0.0	85974 Kidney NAT (OD04450-03)	0.0
98861 Colon cancer (OD06297-04)	0.0	Kidney Cancer Clontech 8120613	0.0
98862 Colon cancer NAT (OD06297-015)	0.0	Kidney NAT Clontech 8120614	0.0
83237 CC Gr.2 ascend colon (ODO3921)	0.0	Kidney Cancer Clontech 9010320	0.0
83238 CC NAT (ODO3921)	0.0	Kidney NAT Clontech 9010321	0.0
97766 Colon cancer metastasis (OD06104)	0.0	Kidney Cancer Clontech 8120607	0.0
97767 Lung NAT (OD06104)	0.0	Kidney NAT Clontech 8120608	0.0
87472 Colon mets to lung (OD04451-01)	0.0	Normal Uterus GENPAK 061018	16.6
87473 Lung NAT (OD04451-02)	0.0	Uterus Cancer GENPAK 064011	0.0
Normal Prostate Clontech A+ 6546-1 (8090438)	0.0	Normal Thyroid Clontech A+ 6570-1 (7080817)	0.0
84140 Prostate Cancer (OD04410)	0.0	Thyroid Cancer GENPAK 064010	0.0
84141 Prostate NAT (OD04410)	0.0	Thyroid Cancer INVITROGEN A302152	0.0
Normal Ovary Res. Gen.	0.0	Thyroid NAT INVITROGEN A302153	0.0

98863 Ovarian cancer (OD06283-03)	0.0	Normal Breast GENPAK 061019	0.0
98865 Ovarian cancer NAT/fallopian tube (OD06283-07)	0.0	84877 Breast Cancer (OD04566)	0.0
Ovarian Cancer GENPAK 064008	**100.0**	Breast Cancer Res. Gen. 1024	0.0
97773 Ovarian cancer (OD06145)	0.0	85975 Breast Cancer (OD04590-01)	0.0
97775 Ovarian cancer NAT (OD06145)	0.0	85976 Breast Cancer Mets (OD04590-03)	0.0
98853 Ovarian cancer (OD06455-03)	0.0	87070 Breast Cancer Metastasis (OD04655-05)	0.0
98854 Ovarian NAT (OD06455-07) Fallopian tube	0.0	GENPAK Breast Cancer 064006	0.0
Normal Lung GENPAK 061010	0.0	Breast Cancer Clontech 9100266	0.0
92337 Invasive poor diff. lung adeno (ODO4945-01	17.1	Breast NAT Clontech 9100265	0.0
92338 Lung NAT (ODO4945-03)	0.0	Breast Cancer INVITROGEN A209073	0.0
84136 Lung Malignant Cancer (OD03126)	0.0	Breast NAT INVITROGEN A2090734	0.0
84137 Lung NAT (OD03126)	0.0	97763 Breast cancer (OD06083)	33.2
90372 Lung Cancer (OD05014A)	0.0	97764 Breast cancer node metastasis (OD06083)	0.0
90373 Lung NAT (OD05014B)	0.0	Normal Liver GENPAK 061009	0.0
97761 Lung cancer (OD06081)	0.0	Liver Cancer Research Genetics RNA 1026	0.0
97762 Lung cancer NAT (OD06081)	0.0	Liver Cancer Research Genetics RNA 1025	0.0
85950 Lung Cancer (OD04237-01)	0.0	Paired Liver Cancer Tissue Research Genetics RNA 6004-T	0.0
85970 Lung NAT (OD04237-02)	0.0	Paired Liver Tissue Research Genetics RNA 6004-N	0.0
83255 Ocular Mel Met to Liver (ODO4310)	0.0	Paired Liver Cancer Tissue Research Genetics RNA 6005-T	0.0
83256 Liver NAT (ODO4310)	0.0	Paired Liver Tissue Research Genetics RNA 6005-N	0.0
84139 Melanoma Mets to Lung (OD04321)	0.0	Liver Cancer GENPAK 064003	72.5
84138 Lung NAT (OD04321)	0.0	Normal Bladder GENPAK 061001	0.0
Normal Kidney GENPAK 061008	0.0	Bladder Cancer Research Genetics RNA 1023	0.0
83786 Kidney Ca, Nuclear grade 2 (OD04338)	0.0	Bladder Cancer INVITROGEN A302173	0.0
83787 Kidney NAT (OD04338)	0.0	Normal Stomach GENPAK 061017	0.0

83788 Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	Gastric Cancer Clontech 9060397	0.0
83789 Kidney NAT (OD04339)	0.0	NAT Stomach Clontech 9060396	0.0
83790 Kidney Ca, Clear cell type (OD04340)	0.0	Gastric Cancer Clontech 9060395	17.0
83791 Kidney NAT (OD04340)	0.0	NAT Stomach Clontech 9060394	0.0
83792 Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer GENPAK 064005	0.0

Table 23. Panel 4D

Tissue Name	Relative Expression(%) 4dx4tm5519t_ag1639_b2	Tissue Name	Relative Expression(%) 4dx4tm5519t_ag1639_b2
93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.4	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.0	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting day 4-6 in IL-2	0.0	93781_HUVEC (Endothelial)_IL-11	0.0
93571_Secondary Tr1_resting day 4-6 in IL-2	0.0	93583_Lung Microvascular Endothelial Cells_none	0.0
93568_primary Th1_anti-CD28/anti-CD3	0.0	93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93569_primary Th2_anti-CD28/anti-CD3	0.0	92662_Microvascular Dermal endothelium_none	0.0
93570_primary Tr1_anti-CD28/anti-CD3	0.0	92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93565_primary Th1_resting dy 4-6 in IL-2	0.0	93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0	93347_Small Airway Epithelium_none	0.0
93567_primary Tr1_resting dy 4-6 in IL-2	0.0	93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3	0.0	92668_Coronery Artery SMC_resting	0.0
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-	0.0	92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b	0.0

CD3		(1 ng/ml)	
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0	93107_astrocytes_resting	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.0	93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	0.0	92666_KU-812 (Basophil)_resting	0.0
93354_CD4_none	0.0	92667_KU-812 (Basophil)_PMA/ionoycin	0.7
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	0.0	93579_CCD1106 (Keratinocytes)_none	0.0
93103_LAK cells_resting	0.0	93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.0
93788_LAK cells_IL-2	0.0	93791_Liver Cirrhosis	9.1
93787_LAK cells_IL-2+IL-12	0.0	93792_Lupus Kidney	0.3
93789_LAK cells_IL-2+IFN gamma	0.0	93577_NCI-H292	0.0
93790_LAK cells_IL-2+ IL-18	0.0	93358_NCI-H292_IL-4	0.0
93104_LAK cells_PMA/ionomycin and IL-18	0.0	93360_NCI-H292_IL-9	0.0
93578_NK Cells IL-2_resting	0.0	93359_NCI-H292_IL-13	0.0
93109_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93357_NCI-H292_IFN gamma	0.0
93110_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93777_HPAEC_-	0.0
93111_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93778_HPAEC_IL-1 beta/TNA alpha	0.0
93112_Mononuclear Cells (PBMCs)_resting	0.0	93254_Normal Human Lung Fibroblast_none	0.0
93113_Mononuclear Cells (PBMCs)_PWM	0.0	93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.0
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0	93257_Normal Human Lung Fibroblast_IL-4	0.3
93249_Ramos (B cell)_none	0.0	93256_Normal Human Lung Fibroblast_IL-9	0.0
93250_Ramos (B cell)_ionomycin	0.0	93255_Normal Human Lung Fibroblast_IL-13	0.0
93349_B lymphocytes_PWM	0.0	93258_Normal Human Lung Fibroblast_IFN gamma	0.0
93350_B lymphoytes_CD40L and IL-4	0.0	93106_Dermal Fibroblasts CCD1070_resting	0.0
92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.0	93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	0.0
93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	0.0

93356_Dendritic Cells_none	1.1	93772_dermal fibroblast_IFN gamma	0.0
93355_Dendritic Cells_LPS 100 ng/ml	0.0	93771_dermal fibroblast_IL-4	0.0
93775_Dendritic Cells_anti-CD40	1.7	93259_IBD Colitis 1**	**100.0**
93774_Monocytes_resting	0.0	93260_IBD Colitis 2	2.4
93776_Monocytes_LPS 50 ng/ml	0.0	93261_IBD Crohns	0.0
93581_Macrophages_resting	0.4	735010_Colon_normal	5.6
93582_Macrophages_LPS 100 ng/ml	0.0	735019_Lung_none	0.0
93098_HUVEC (Endothelial)_none	0.0	64028-1_Thymus_none	0.0
93099_HUVEC (Endothelial)_starved	0.0	64030-1_Kidney_none	0.0

Panel 1.3D Summary: Expression of the AP001804_C gene in this panel is highest in the spleen. Expression is detected at a much lower level in a melanoma and a breast cancer cell line. This profile may indicate that the expression of this gene is restricted to splenic lymphoid tissues and thus may be useful as a marker of this tissue.

Panel 2.2 Summary: Significant expression of the AP001804_C gene on panel 2.2 is restricted to one ovarian cancer and one liver cancer. This information suggests that this gene may be of use in the diagnosis and/or treatment of ovarian or liver cancer.

Panel 4D Summary: The AP001804_C transcript is expressed in colitis 1, colitis 2, an activated basophil cell line and in dendritic cells. The protein encoded for by this antigen may be important in the inflammatory process and particularly in the function of activated dendritic cells or basophils. Antagonistic antibodies or small molecule therapeutics against the AP001804_C protein may therefore reduce or inhibit inflammation in the bowel due to IBD by specifically targeting dendritic cells and basophils or other related cell types. This gene was found to be expressed in spleen in Panel 1.3D.

D. GPCR4 (also known as AP001804_D or CG54353-01)

Expression of gene AP001804_D was assessed using the primer-probe set Ag3091, described in Table 24. Results of the RTQ-PCR runs are shown in Tables 25, 26, and 27.

Table 24. Probe Name Ag3091

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-GGTGCATGACTCAGCTGTTT-3'	58.9	20	287	89

Probe	FAM-5'-TCATCTCTGAATGTTACATGTTGACCTCA-3'-TAMRA	65.9	29	323	90
Reverse	5'-GCCACATAGCGATCATATGC-3'	59.1	20	355	91

Table 25. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5823f_ag3091_b1	Tissue Name	Relative Expression(%) 1.3dx4tm5823f_ag3091_b1
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	0.0	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	0.0	Lung (fetal)	1.6
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	0.0	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	8.1	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	0.0
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	2.2
CNS ca. (glio/astro) U-118-MG	4.7	Lung ca. (non-s.cell) NCI-H23	0.0
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	0.0
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	0.0
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.0	Mammary gland	0.0
CNS ca. (glio) U251	0.0	Breast ca.* (pl. effusion) MCF-7	74.9
CNS ca. (glio) SF-295	1.2	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	**100.0**
Heart	0.0	Breast ca. BT-549	0.0
Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.0	Ovarian ca. OVCAR-4	15.7
Spleen	0.0	Ovarian ca. OVCAR-5	0.0

Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	5.5	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	8.3
Colon ca. HCT-116	0.0	Testis	16.3
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	20.5
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table 26. Panel 2.2

Tissue Name	Relative Expression(%) 2.2x4tm6415f_ ag3091_b1	Tissue Name	Relative Expression(%) 2.2x4tm6415f_ ag3091_b1
Normal Colon GENPAK 061003	0.0	83793 Kidney NAT (OD04348)	0.0
97759 Colon cancer (OD06064)	0.0	98938 Kidney malignant cancer (OD06204B)	0.0
97760 Colon cancer NAT (OD06064)	0.0	98939 Kidney normal adjacent tissue (OD06204E)	10.7
97778 Colon cancer (OD06159)	0.0	85973 Kidney Cancer (OD04450-01)	0.0
97779 Colon cancer NAT (OD06159)	**0.0**	85974 Kidney NAT (OD04450-03)	0.0
98861 Colon cancer (OD06297-04)	0.0	Kidney Cancer Clontech 8120613	0.0
98862 Colon cancer NAT (OD06297-015)	0.0	Kidney NAT Clontech 8120614	0.0
83237 CC Gr.2 ascend colon (ODO3921)	0.0	Kidney Cancer Clontech 9010320	0.0
83238 CC NAT (ODO3921)	0.0	Kidney NAT Clontech 9010321	15.6
97766 Colon cancer metastasis (OD06104)	0.0	Kidney Cancer Clontech 8120607	0.0
97767 Lung NAT (OD06104)	8.7	Kidney NAT Clontech 8120608	0.0
87472 Colon mets to lung (OD04451-01)	0.0	Normal Uterus GENPAK 061018	13.7
87473 Lung NAT (OD04451-02)	0.0	Uterus Cancer GENPAK 064011	0.0
Normal Prostate Clontech A+ 6546-1 (8090438)	0.0	Normal Thyroid Clontech A+ 6570-1 (7080817)	0.0

84140 Prostate Cancer (OD04410)	0.0
84141 Prostate NAT (OD04410)	0.0
Normal Ovary Res. Gen.	0.0
98863 Ovarian cancer (OD06283-03)	0.0
98865 Ovarian cancer NAT/fallopian tube (OD06283-07)	0.0
Ovarian Cancer GENPAK 064008	4.1
97773 Ovarian cancer (OD06145)	0.0
97775 Ovarian cancer NAT (OD06145)	22.8
98853 Ovarian cancer (OD06455-03)	0.0
98854 Ovarian NAT (OD06455-07) Fallopian tube	0.0
Normal Lung GENPAK 061010	0.0
92337 Invasive poor diff. lung adeno (ODO4945-01	11.2
92338 Lung NAT (ODO4945-03)	0.0
84136 Lung Malignant Cancer (OD03126)	0.0
84137 Lung NAT (OD03126)	0.0
90372 Lung Cancer (OD05014A)	0.0
90373 Lung NAT (OD05014B)	4.0
97761 Lung cancer (OD06081)	0.0
97762 Lung cancer NAT (OD06081)	0.0
85950 Lung Cancer (OD04237-01)	0.0
85970 Lung NAT (OD04237-02)	0.0
83255 Ocular Mel Met to Liver (ODO4310)	0.0
83256 Liver NAT (ODO4310)	0.0
84139 Melanoma Mets to Lung (OD04321)	0.0
84138 Lung NAT (OD04321)	0.0
Thyroid Cancer GENPAK 064010	0.0
Thyroid Cancer INVITROGEN A302152	0.0
Thyroid NAT INVITROGEN A302153	0.0
Normal Breast GENPAK 061019	0.0
84877 Breast Cancer (OD04566)	0.0
Breast Cancer Res. Gen. 1024	0.0
85975 Breast Cancer (OD04590-01)	0.0
85976 Breast Cancer Mets (OD04590-03)	0.0
87070 Breast Cancer Metastasis (OD04655-05)	0.0
GENPAK Breast Cancer 064006	0.0
Breast Cancer Clontech 9100266	0.0
Breast NAT Clontech 9100265	0.0
Breast Cancer INVITROGEN A209073	0.0
Breast NAT INVITROGEN A2090734	0.0
97763 Breast cancer (OD06083)	0.0
97764 Breast cancer node metastasis (OD06083)	0.0
Normal Liver GENPAK 061009	0.0
Liver Cancer Research Genetics RNA 1026	0.0
Liver Cancer Research Genetics RNA 1025	0.0
Paired Liver Cancer Tissue Research Genetics RNA 6004-T	0.0
Paired Liver Tissue Research Genetics RNA 6004-N	0.0
Paired Liver Cancer Tissue Research Genetics RNA 6005-T	0.0
Paired Liver Tissue Research Genetics RNA 6005-N	0.0
Liver Cancer GENPAK 064003	14.5
Normal Bladder GENPAK 061001	0.0

Normal Kidney GENPAK 061008	0.0	Bladder Cancer Research Genetics RNA 1023	0.0
83786 Kidney Ca, Nuclear grade 2 (OD04338)	13.0	Bladder Cancer INVITROGEN A302173	**100.0**
83787 Kidney NAT (OD04338)	14.8	Normal Stomach GENPAK 061017	0.0
83788 Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	Gastric Cancer Clontech 9060397	0.0
83789 Kidney NAT (OD04339)	0.0	NAT Stomach Clontech 9060396	5.4
83790 Kidney Ca, Clear cell type (OD04340)	0.0	Gastric Cancer Clontech 9060395	15.0
83791 Kidney NAT (OD04340)	0.0	NAT Stomach Clontech 9060394	0.0
83792 Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer GENPAK 064005	0.0

Table 27. Panel 4D

Tissue Name	Relative Expression(%) 4dx4tm5055f_ag3091_b2	Tissue Name	Relative Expression(%) 4dx4tm5055f_ag3091_b2
93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.0	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.0	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting day 4-6 in IL-2	0.0	93781_HUVEC (Endothelial)_IL-11	0.0
93571_Secondary Tr1_resting day 4-6 in IL-2	0.0	93583_Lung Microvascular Endothelial Cells_none	0.0
93568_primary Th1_anti-CD28/anti-CD3	0.0	93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93569_primary Th2_anti-CD28/anti-CD3	0.0	92662_Microvascular Dermal endothelium_none	0.0
93570_primary Tr1_anti-CD28/anti-CD3	0.0	92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93565_primary Th1_resting dy 4-6 in IL-2	0.0	93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0	93347_Small Airway Epithelium_none	0.0
93567_primary Tr1_resting dy 4-6 in IL-2	0.0	93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93351_CD45RA CD4	0.0	92668_Coronery Artery	0.0

lymphocyte_anti-CD28/anti-CD3		SMC_resting	
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3	0.0	92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0	93107_astrocytes_resting	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.0	93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	0.0	92666_KU-812 (Basophil)_resting	0.0
93354_CD4_none	0.0	92667_KU-812 (Basophil)_PMA/ionoycin	0.0
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	0.0	93579_CCD1106 (Keratinocytes)_none	0.0
93103_LAK cells_resting	0.0	93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.0
93788_LAK cells_IL-2	0.0	93791_Liver Cirrhosis	0.0
93787_LAK cells_IL-2+IL-12	0.0	93792_Lupus Kidney	0.0
93789_LAK cells_IL-2+IFN gamma	0.0	93577_NCI-H292	0.0
93790_LAK cells_IL-2+ IL-18	0.0	93358_NCI-H292_IL-4	0.0
93104_LAK cells_PMA/ionomycin and IL-18	0.0	93360_NCI-H292_IL-9	0.0
93578_NK Cells IL-2_resting	0.0	93359_NCI-H292_IL-13	0.0
93109_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93357_NCI-H292_IFN gamma	0.0
93110_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93777_HPAEC_-	0.0
93111_Mixed Lymphocyte Reaction_Two Way MLR	0.0	93778_HPAEC_IL-1 beta/TNA alpha	0.0
93112_Mononuclear Cells (PBMCs)_resting	0.0	93254_Normal Human Lung Fibroblast_none	0.0
93113_Mononuclear Cells (PBMCs)_PWM	0.0	93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.0
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0	93257_Normal Human Lung Fibroblast_IL-4	17.0
93249_Ramos (B cell)_none	0.0	93256_Normal Human Lung Fibroblast_IL-9	0.0
93250_Ramos (B cell)_ionomycin	0.0	93255_Normal Human Lung Fibroblast_IL-13	0.0
93349_B lymphocytes_PWM	0.0	93258_Normal Human Lung Fibroblast_IFN gamma	0.0
93350_B lymphoytes_CD40L and IL-4	0.0	93106_Dermal Fibroblasts CCD1070_resting	0.0
92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.0	93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	0.0

93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	0.0
93356_Dendritic Cells_none	31.7	93772_dermal fibroblast_IFN gamma	0.0
93355_Dendritic Cells_LPS 100 ng/ml	4.6	93771_dermal fibroblast_IL-4	0.0
93775_Dendritic Cells_anti-CD40	43.3	93259_IBD Colitis 1**	0.0
93774_Monocytes_resting	0.0	93260_IBD Colitis 2	2.9
93776_Monocytes_LPS 50 ng/ml	0.0	93261_IBD Crohns	7.7
93581_Macrophages_resting	**100.0**	735010_Colon_normal	40.3
93582_Macrophages_LPS 100 ng/ml	0.0	735019_Lung_none	55.9
93098_HUVEC (Endothelial)_none	0.0	64028-1_Thymus_none	0.0
93099_HUVEC (Endothelial)_starved	0.0	64030-1_Kidney_none	0.0

Panel 1.3D Summary: The expression of the AP001804_D gene appears to be restricted to two breast cancer cell lines. Interestingly both of these cell lines are positive for estrogen receptor expression. Thus, this gene may be a marker for estrogen receptor positive breast cancer cells. Further, therapeutic modulation of this gene may be of use in the treatment of breast cancer or other breast related disease.

Panel 2.2 Summary: Two RTQ-PCR experiments were performed using Ag3091. In one experiment, AP001804_D gene expression was low to undetectable (CT values >35) in all samples. In the other experiment, expression was low/undectable in all samples except a single bladder cancer cell line (CT=34.5). Expression levels are too low for reliable analysis.

Panel 4D Summary: The AP001804_D transcript is detectable in resting macrophages and not at significant levels in other cell types. Antibody or protein therapeutics designed against the AP001804_D protein encoded for by this transcript could reduce or inhibit inflammation in asthma, emphysema, allergy, psoriasis, arthritis, or any other condition in which macrophage localization/activation is important.

E. GPCR5 (also known as AP001804_E or CG54362-01)

Expression of gene AP001804_E was assessed using the primer-probe sets Ag2359, Ag2358, and Ag1640 (identical sequences), described in Table 28. Results of the RTQ-PCR runs are shown in Tables 29 and 30.

Table 28. Probe Name Ag2359/Ag2358/Ag1640

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-CCATGTCAGTGAGCTGGTATTT-3'	59.1	22	574	92
Probe	FAM-5'-TGGAGTAATCACCATGCTATCCAGCA-3'-TAMRA	67.7	26	607	93
Reverse	5'-TCAAAGCGTAAGAGATGACGAT-3'	59	22	638	94

Table 29. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5396f _ag1640_a2	Tissue Name	Relative Expression(%) 1.3dx4tm5396f _ag1640_a2
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	39.2
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	0.0	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	0.0	Lung (fetal)	0.0
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	38.8	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	0.0	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	0.0
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
CNS ca. (glio/astro) U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	0.0
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	0.0
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	0.0
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.0	Mammary gland	0.0
CNS ca. (glio) U251	20.2	Breast ca.* (pl. effusion) MCF-7	87.4
CNS ca. (glio) SF-295	21.0	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	76.7

Heart	21.3	Breast ca. BT-549	0.0
Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.0	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	0.0
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	0.0	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	44.8
Colon ca. HCT-116	0.0	Testis	16.4
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	**100.0**
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table 30. Panel 2D

Tissue Name	Relative Expression(%) 2dx4tm4937f_ ag2359_a1	Relative Expression(%) 2dx4tm4923f_ ag2358_b2
Normal Colon GENPAK 061003	2.3	8.0
83219 CC Well to Mod Diff (ODO3866)	4.2	4.4
83220 CC NAT (ODO3866)	9.2	1.7
83221 CC Gr.2 rectosigmoid (ODO3868)	0.0	0.0
83222 CC NAT (ODO3868)	0.0	0.0
83235 CC Mod Diff (ODO3920)	0.0	0.0
83236 CC NAT (ODO3920)	0.0	0.0
83237 CC Gr.2 ascend colon (ODO3921)	0.0	0.0
83238 CC NAT (ODO3921)	1.2	1.1
83241 CC from Partial Hepatectomy (ODO4309)	0.0	0.0
83242 Liver NAT (ODO4309)	0.0	0.0
87472 Colon mets to lung (OD04451-01)	0.0	4.5
87473 Lung NAT (OD04451-02)	0.0	0.0
Normal Prostate Clontech A+ 6546-1	0.0	0.0
84140 Prostate Cancer (OD04410)	0.0	0.0
84141 Prostate NAT (OD04410)	0.0	0.0

87073 Prostate Cancer (OD04720-01)	0.0	0.0
87074 Prostate NAT (OD04720-02)	0.0	0.0
Normal Lung GENPAK 061010	0.0	0.0
83239 Lung Met to Muscle (ODO4286)	2.1	0.0
83240 Muscle NAT (ODO4286)	0.0	0.0
84136 Lung Malignant Cancer (OD03126)	0.0	1.3
84137 Lung NAT (OD03126)	0.0	0.9
84871 Lung Cancer (OD04404)	0.0	0.0
84872 Lung NAT (OD04404)	0.0	0.0
84875 Lung Cancer (OD04565)	0.0	0.0
84876 Lung NAT (OD04565)	0.0	0.0
85950 Lung Cancer (OD04237-01)	0.0	1.7
85970 Lung NAT (OD04237-02)	0.0	0.0
83255 Ocular Mel Met to Liver (ODO4310)	1.8	0.0
83256 Liver NAT (ODO4310)	0.0	0.0
84139 Melanoma Mets to Lung (OD04321)	0.0	0.0
84138 Lung NAT (OD04321)	0.0	0.0
Normal Kidney GENPAK 061008	3.4	5.1
83786 Kidney Ca, Nuclear grade 2 (OD04338)	8.1	13.4
83787 Kidney NAT (OD04338)	0.0	4.1
83788 Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	5.6
83789 Kidney NAT (OD04339)	1.6	0.0
83790 Kidney Ca, Clear cell type (OD04340)	0.0	0.0
83791 Kidney NAT (OD04340)	3.7	0.0
83792 Kidney Ca, Nuclear grade 3 (OD04348)	0.0	0.0
83793 Kidney NAT (OD04348)	0.0	0.0
87474 Kidney Cancer (OD04622-01)	0.0	0.0
87475 Kidney NAT (OD04622-03)	0.0	0.0
85973 Kidney Cancer (OD04450-01)	0.0	0.0
85974 Kidney NAT (OD04450-03)	0.0	0.0
Kidney Cancer Clontech 8120607	0.0	0.0
Kidney NAT Clontech 8120608	0.0	0.0
Kidney Cancer Clontech 8120613	0.0	0.0
Kidney NAT Clontech 8120614	0.0	0.0
Kidney Cancer Clontech 9010320	0.0	0.0
Kidney NAT Clontech 9010321	0.0	0.0
Normal Uterus GENPAK 061018	0.0	0.8
Uterus Cancer GENPAK 064011	0.0	0.0
Normal Thyroid Clontech A+ 6570-1	0.0	0.0
Thyroid Cancer GENPAK 064010	0.0	0.0
Thyroid Cancer INVITROGEN A302152	0.0	0.0
Thyroid NAT INVITROGEN A302153	0.0	0.0
Normal Breast GENPAK 061019	0.0	0.0
84877 Breast Cancer (OD04566)	0.0	0.0
85975 Breast Cancer (OD04590-01)	0.0	0.0
85976 Breast Cancer Mets (OD04590-03)	2.2	0.0

87070 Breast Cancer Metastasis (OD04655-05)	0.0	0.0
GENPAK Breast Cancer 064006	1.7	0.0
Breast Cancer Res. Gen. 1024	0.0	3.3
Breast Cancer Clontech 9100266	0.0	0.0
Breast NAT Clontech 9100265	0.0	0.0
Breast Cancer INVITROGEN A209073	0.0	1.8
Breast NAT INVITROGEN A2090734	0.0	0.0
Normal Liver GENPAK 061009	0.0	0.0
Liver Cancer GENPAK 064003	21.7	6.5
Liver Cancer Research Genetics RNA 1025	0.0	0.0
Liver Cancer Research Genetics RNA 1026	0.0	0.0
Paired Liver Cancer Tissue Research Genetics RNA 6004-T	1.4	0.0
Paired Liver Tissue Research Genetics RNA 6004-N	0.0	0.0
Paired Liver Cancer Tissue Research Genetics RNA 6005-T	0.0	0.0
Paired Liver Tissue Research Genetics RNA 6005-N	0.0	0.0
Normal Bladder GENPAK 061001	0.0	0.0
Bladder Cancer Research Genetics RNA 1023	0.0	0.0
Bladder Cancer INVITROGEN A302173	45.9	57.8
87071 Bladder Cancer (OD04718-01)	0.0	0.0
87072 Bladder Normal Adjacent (OD04718-03)	0.0	0.0
Normal Ovary Res. Gen.	0.0	0.0
Ovarian Cancer GENPAK 064008	0.0	0.0
87492 Ovary Cancer (OD04768-07)	**100.0**	**100.0**
87493 Ovary NAT (OD04768-08)	0.0	0.0
Normal Stomach GENPAK 061017	3.5	0.0
Gastric Cancer Clontech 9060358	0.0	0.0
NAT Stomach Clontech 9060359	0.0	0.0
Gastric Cancer Clontech 9060395	0.0	0.0
NAT Stomach Clontech 9060394	0.0	0.0
Gastric Cancer Clontech 9060397	0.0	0.0
NAT Stomach Clontech 9060396	0.0	0.0
Gastric Cancer GENPAK 064005	0.0	1.8

Panel 1.3 D Summary: Ag1640 Significant expression of the AP001804_E gene is restricted to one melanoma cell line indicating that this gene may be a useful marker for melanoma. Ag2358/Ag2359 Expression of the AP001804_E gene was low/undetectable (CT values > 35) in all samples on this panel.

Panel 2D Summary: Ag2359/Ag2358 The AP001804_E gene is most abundantly expressed in a sample of ovarian cancer with limited, very low level of expression in other tissues. Thus, this gene may be useful in distinguishing ovarian cancers from other tissues. Therapeutic modulation of this gene may also be useful in the treatment of ovarian cancers.

Panel 2.2 Summary: Ag1640 Expression of the AP001804_E gene was low/undetectable (CT values > 35) in all samples on this panel and thus the results not shown.

Panel 4D Summary: Ag2359/Ag2358/Ag1640 Expression of the AP001804_E gene was low/undetectable (CT values > 35) in all samples on this panel and thus the results not shown.

F. GPCR6 (also known as AP000868_A or CG54263-01)

Expression of gene AP000868_A was assessed using the primer-probe set Ag1629, described in Table 31. Results of the RTQ-PCR run is shown in Table 32.

Table 31. Probe Name Ag1629

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-AAATCTGGTACACCACCACAGT-3'	58.4	22	203	95
Probe	FAM-5'-CATCCCCAAACTGCTAGGAACCTTTG-3'-TAMRA	68.5	26	225	96
Reverse	5'-AGCAGGACATGCAGATTACTGT-3'	58.9	22	262	97

Table 32. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3dx4tm5395f_ag1629_b2	Tissue Name	Relative Expression(%) 1.3dx4tm5395f_ag1629_b2
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	21.2	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	5.7	Liver (fetal)	0.0
Brain (amygdala)	78.2	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	22.7	Lung	0.0
Brain (hippocampus)	**100.0**	Lung (fetal)	0.0
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0

Brain (thalamus)	0.0	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	0.0	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell)NCI-H460	0.0
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
CNS ca. (glio/astro) U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	4.1
CNS ca. (astro) SW1783	0.0	Lung ca (non-s.cell) HOP-62	0.0
CNS ca.* (neuro; met) SK-N-AS	0.0	Lung ca. (non-s.cl) NCI-H522	0.0
CNS ca. (astro) SF-539	0.0	Lung ca. (squam.) SW 900	0.0
CNS ca. (astro) SNB-75	0.0	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.0	Mammary gland	0.0
CNS ca. (glio) U251	11.1	Breast ca.* (pl. effusion) MCF-7	0.0
CNS ca. (glio) SF-295	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl. effusion) T47D	0.0
Heart	0.0	Breast ca. BT-549	0.0
Fetal Skeletal	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	0.0
Thymus	0.0	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	0.0
Lymph node	3.3	Ovarian ca. OVCAR-8	0.0
Colorectal	0.0	Ovarian ca. IGROV-1	0.0
Stomach	2.5	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* (SW480 met)SW620	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	0.0
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
83219 CC Well to Mod Diff (ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.0	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0

Trachea	0.0	Melanoma* (met) SK-MEL-5	0.0
Kidney	0.0	Adipose	9.1

Panel 1.3D Summary: The AP000868_A transcript appears to be brain specific (or at least to show highly preferential expression in brain), especially in the hippocampus and amygdala. These regions are of great interest as both have been implicated in Alzheimer's disease, schizophrenia, and bipolar disorder. Furthermore, the hippocampus is critical in the development of long-term memories, and the amygdala is involved in the processing of emotion (e.g., fear, etc). Because this transcript encodes for a GPCR, the AP000868_A protein is also a potential small molecule target for the treatment/prevention of both neurodegenerative and psychiatric disorders. In addition, the AP000868_A gene product could possibly be targeted in normal, healthy populations for modulation of memory and fear/anxiety.

Panel 2.2 Summary: Expression of the AP000868_A gene was low/undetectable (CT values > 35) in all samples on this panel and thus the results not shown.

Panel 4D Summary: Expression of the AP000868_A gene was low/undetectable (CT values > 35) in all samples on this panel and thus the results not shown.

Panel CNSD.01 Summary: Expression of the AP000868_A gene was low/undetectable (CT values > 35) in all samples on this panel and thus the results not shown.

G. GPCR7 (also known as 20722608_EXT or CG51505-01)

Expression of gene 20722608_EXT was assessed using the primer-probe set Ag1629, described in Table 33. Results of the RTQ-PCR run is shown in Table 34.

Table 33. Probe Name Gpcr27

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-TGCCCTTGTCATCCTGAGC-3'		19	187	98
Probe	FAM-5'-CCTGGCATTCACTACTGGGTCCTACGTGTA-3'-TAMRA		30	208	99
Reverse	5'-GGGATACGCAGGATGGTAGAAA-3'		22	243	100

Table 34. Panel 1

Tissue Name	Relative Expression(%) 1xtm240f_ gpcr27	Tissue Name	Relative Expression(%) 1xtm240f_ gpcr27
Endothelial cells	2.6	Kidney (fetal)	2.5
Endothelial cells (treated)	5.2	Renal ca. 786-0	0
Pancreas	**100**	Renal ca. A498	0.6
Pancreatic ca. CAPAN 2	0	Renal ca. RXF 393	0.7
Adipose	10.4	Renal ca. ACHN	0
Adrenal gland	0	Renal ca. UO-31	2.4
Thyroid	0.7	Renal ca. TK-10	0
Salavary gland	0	Liver	0
Pituitary gland	0	Liver (fetal)	0
Brain (fetal)	3.3	Liver ca. (hepatoblast) HepG2	0
Brain (whole)	5.4	Lung	0
Brain (amygdala)	2.2	Lung (fetal)	0
Brain (cerebellum)	1.2	Lung ca. (small cell) LX-1	0
Brain (hippocampus)	3	Lung ca. (small cell) NCI-H69	11.7
Brain (substantia nigra)	0.6	Lung ca. (s.cell var.) SHP-77	0
Brain (thalamus)	0	Lung ca. (large cell)NCI-H460	0
Brain (hypothalamus)	0.2	Lung ca. (non-sm. cell) A549	0
Spinal cord	0	Lung ca. (non-s.cell) NCI-H23	0
CNS ca. (glio/astro) U87-MG	1.5	Lung ca (non-s.cell) HOP-62	0.6
CNS ca. (glio/astro) U-118-MG	0	Lung ca. (non-s.cl) NCI-H522	0
CNS ca. (astro) SW1783	1.5	Lung ca. (squam.) SW 900	1.3
CNS ca.* (neuro; met) SK-N-AS	0	Lung ca. (squam.) NCI-H596	1.4
CNS ca. (astro) SF-539	0	Mammary gland	1.4
CNS ca. (astro) SNB-75	0	Breast ca.* (pl. effusion) MCF-7	0
CNS ca. (glio) SNB-19	1.2	Breast ca.* (pl.ef) MDA-MB-231	0
CNS ca. (glio) U251	0.6	Breast ca.* (pl. effusion) T47D	21.9
CNS ca. (glio) SF-295	0	Breast ca. BT-549	0
Heart	0	Breast ca. MDA-N	1.5
Skeletal muscle	0	Ovary	0
Bone marrow	0	Ovarian ca. OVCAR-3	0
Thymus	0.6	Ovarian ca. OVCAR-4	1.5
Spleen	1.5	Ovarian ca. OVCAR-5	5.6
Lymph node	0	Ovarian ca. OVCAR-8	0.7
Colon (ascending)	57	Ovarian ca. IGROV-1	2
Stomach	0	Ovarian ca.* (ascites) SK-OV-3	0
Small intestine	0	Uterus	2.2
Colon ca. SW480	0	Placenta	4.1
Colon ca.* (SW480 met)SW620	0	Prostate	0

Colon ca. HT29	0.8	Prostate ca.* (bone met)PC-3	0
Colon ca. HCT-116	0	Testis	99.3
Colon ca. CaCo-2	0.8	Melanoma Hs688(A).T	0
Colon ca. HCT-15	2.9	Melanoma* (met) Hs688(B).T	3.4
Colon ca. HCC-2998	0	Melanoma UACC-62	0
Gastric ca.* (liver met) NCI-N87	2.3	Melanoma M14	10.6
Bladder	0	Melanoma LOX IMVI	1.3
Trachea	1.2	Melanoma* (met) SK-MEL-5	0.2
Kidney	1	Melanoma SK-MEL-28	0.3

Panel 1 Summary: The 20722608_EXT gene is expressed most abundantly in testis, colon and pancreas. Expression in the testis may be due to genomic DNA contamination. The expression of 20722608_EXT gene seems to be specific for pancreas and colon tissues. These tissues both play an important role in the process of digestion and thus, therapeutic modulation of the 20722608_EXT gene may be of utility in the treatment of gastrointestinal disease related to the colon and/or pancreas. In addition, although the 20722608_EXT gene is most highly expressed in the pancreas, it is absent in a pancreatic cancer cell lines suggesting that this gene could be useful in the diagnosis/treatment of pancreatic cancer. 20722608_EXT gene may be involved in signal transduction pathways in either the exocrine or endocrine tissues of the pancreas. Thus, this gene may be a drug target for diseases of the pancreas including Types 1 and 2 diabetes and any or all forms of pancreatitis.

H. 21629632_EXT

Expression of gene 21629632_EXT was assessed using the primer-probe set Ag1539, described in Table 35. Results of the RTQ-PCR run is shown in Table 36, 37, 38, 39 and 40.

Table 35. Probe name Ag1539

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-TTTTATGGGACAATCTCCTTCA-3'	58.6	22	745	101
Probe	FAM-5'-TGTACTTCAAACCCAAGGCCAAGGAT-3'-TAMRA	68.4	26	767	102
Reverse	5'-GAACAATGCGACAGTCTTATCC-3'	58.7	22	801	103

Table 36. Panel 1.2

Tissue Name	Relative Expression(%) 1.2tm2212f_ ag1539	Tissue Name	Relative Expression(%) 1.2tm2212f_ ag1539
Endothelial cells	0.1	Renal ca. 786-0	0.7
Endothelial cells (treated)	3.5	Renal ca. A498	3.1
Pancreas	2.7	Renal ca. RXF 393	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. ACHN	3.4
Adrenal Gland (new lot*)	9.5	Renal ca. UO-31	2.2
Thyroid	1.3	Renal ca. TK-10	3.1
Salavary gland	26.1	Liver	2.8
Pituitary gland	2.1	Liver (fetal)	2.6
Brain (fetal)	4.9	Liver ca. (hepatoblast) HepG2	0.5
Brain (whole)	22.8	Lung	0.5
Brain (amygdala)	14.9	Lung (fetal)	0.8
Brain (cerebellum)	14.0	Lung ca. (small cell) LX-1	13.0
Brain (hippocampus)	81.2	Lung ca. (small cell) NCI-H69	2.0
Brain (thalamus)	31.9	Lung ca. (s.cell var.) SHP-77	0.1
Cerebral Cortex	**100.0**	Lung ca. (large cell)NCI-H460	2.8
Spinal cord	3.3	Lung ca. (non-sm. cell) A549	4.1
CNS ca. (glio/astro) U87-MG	1.4	Lung ca. (non-s.cell) NCI-H23	1.2
CNS ca. (glio/astro) U-118-MG	0.3	Lung ca (non-s.cell) HOP-62	8.4
CNS ca. (astro) SW1783	0.4	Lung ca. (non-s.cl) NCI-H522	23.3
CNS ca.* (neuro; met) SK-N-AS	1.7	Lung ca. (squam.) SW 900	13.8
CNS ca. (astro) SF-539	1.7	Lung ca. (squam.) NCI-H596	1.3
CNS ca. (astro) SNB-75	1.9	Mammary gland	6.6
CNS ca. (glio) SNB-19	5.0	Breast ca.* (pl. effusion) MCF-7	1.2
CNS ca. (glio) U251	3.1	Breast ca.* (pl.ef) MDA-MB-231	0.5
CNS ca. (glio) SF-295	25.9	Breast ca.* (pl. effusion) T47D	5.4
Heart	46.3	Breast ca. BT-549	37.4
Skeletal Muscle (new lot*)	52.1	Breast ca. MDA-N	1.3
Bone marrow	0.4	Ovary	7.1
Thymus	0.3	Ovarian ca. OVCAR-3	3.7
Spleen	1.2	Ovarian ca. OVCAR-4	1.8
Lymph node	0.6	Ovarian ca. OVCAR-5	27.7
Colorectal	0.2	Ovarian ca. OVCAR-8	6.6
Stomach	2.5	Ovarian ca. IGROV-1	5.7
Small intestine	7.1	Ovarian ca.* (ascites) SK-OV-3	3.4
Colon ca. SW480	0.3	Uterus	3.2
Colon ca.* (SW480 met)SW620	0.9	Placenta	0.4
Colon ca. HT29	1.5	Prostate	20.2
Colon ca. HCT-116	0.9	Prostate ca.* (bone met)PC-3	3.3
Colon ca. CaCo-2	2.3	Testis	1.3

83219 CC Well to Mod Diff (ODO3866)	0.6	Melanoma Hs688(A).T	0.6
Colon ca. HCC-2998	11.9	Melanoma* (met) Hs688(B).T	0.5
Gastric ca.* (liver met) NCI-N87	4.9	Melanoma UACC-62	3.9
Bladder	5.0	Melanoma M14	1.6
Trachea	0.2	Melanoma LOX IMVI	0.0
Kidney	30.4	Melanoma* (met) SK-MEL-5	0.0
Kidney (fetal)	11.5	Adipose	18.0

Table 37. Panel 1.3D

Tissue Name	Relative Expression(%) 1.3Dtm2998f_ ag1539	Tissue Name	Relative Expression(%) 1.3Dtm2998f_ ag1539
Liver adenocarcinoma	1.7	Kidney (fetal)	1.8
Pancreas	0.5	Renal ca. 786-0	1.6
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	2.2
Adrenal gland	2.7	Renal ca. RXF 393	0.5
Thyroid	4.4	Renal ca. ACHN	1.7
Salivary gland	2.0	Renal ca. UO-31	0.0
Pituitary gland	7.4	Renal ca. TK-10	1.2
Brain (fetal)	21.6	Liver	0.2
Brain (whole)	26.6	Liver (fetal)	1.6
Brain (amygdala)	30.8	Liver ca. (hepatoblast) HepG2	0.9
Brain (cerebellum)	7.6	Lung	1.7
Brain (hippocampus)	**100.0**	Lung (fetal)	3.5
Brain (substantia nigra)	5.0	Lung ca. (small cell) LX-1	4.1
Brain (thalamus)	15.8	Lung ca. (small cell) NCI-H69	1.2
Cerebral Cortex	76.8	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	4.3	Lung ca. (large cell)NCI-H460	0.3
CNS ca. (glio/astro) U87-MG	0.0	Lung ca. (non-sm. cell) A549	2.1
CNS ca. (glio/astro) U-118-MG	0.3	Lung ca. (non-s.cell) NCI-H23	0.6
CNS ca. (astro) SW1783	0.7	Lung ca (non-s.cell) HOP-62	2.2
CNS ca.* (neuro; met) SK-N-AS	0.9	Lung ca. (non-s.cl) NCI-H522	4.0
CNS ca. (astro) SF-539	1.5	Lung ca. (squam.) SW 900	2.6
CNS ca. (astro) SNB-75	3.6	Lung ca. (squam.) NCI-H596	0.0
CNS ca. (glio) SNB-19	0.7	Mammary gland	1.8
CNS ca. (glio) U251	3.6	Breast ca.* (pl. effusion) MCF-7	0.0
CNS ca. (glio) SF-295	15.6	Breast ca.* (pl.ef) MDA-MB-231	0.9
Heart (fetal)	6.1	Breast ca.* (pl. effusion) T47D	1.6
Heart	2.4	Breast ca. BT-549	0.8
Fetal Skeletal	70.7	Breast ca. MDA-N	0.0

Skeletal muscle	0.5	Ovary	7.7
Bone marrow	0.0	Ovarian ca. OVCAR-3	1.0
Thymus	1.1	Ovarian ca. OVCAR-4	0.0
Spleen	0.4	Ovarian ca. OVCAR-5	4.8
Lymph node	1.0	Ovarian ca. OVCAR-8	1.8
Colorectal	8.5	Ovarian ca. IGROV-1	1.1
Stomach	2.9	Ovarian ca.* (ascites) SK-OV-3	0.6
Small intestine	4.5	Uterus	4.0
Colon ca. SW480	0.0	Placenta	0.3
Colon ca.* (SW480 met)SW620	0.9	Prostate	4.7
Colon ca. HT29	1.1	Prostate ca.* (bone met)PC-3	2.4
Colon ca. HCT-116	0.1	Testis	5.0
Colon ca. CaCo-2	0.9	Melanoma Hs688(A).T	1.3
83219 CC Well to Mod Diff (ODO3866)	1.2	Melanoma* (met) Hs688(B).T	1.8
Colon ca. HCC-2998	1.8	Melanoma UACC-62	0.7
Gastric ca.* (liver met) NCI-N87	3.3	Melanoma M14	0.3
Bladder	4.2	Melanoma LOX IMVI	0.0
Trachea	2.3	Melanoma* (met) SK-MEL-5	0.4
Kidney	3.3	Adipose	1.1

Table 38. Panel 2D

Tissue Name	Relative Expression(%) 2Dtm2349f_ ag1539	Relative Expression(%) 2dtm2829f_ ag1539
Normal Colon GENPAK 061003	2.2	37.9
83219 CC Well to Mod Diff (ODO3866)	0.2	2.7
83220 CC NAT (ODO3866)	0.2	2.7
83221 CC Gr.2 rectosigmoid (ODO3868)	0.4	7.2
83222 CC NAT (ODO3868)	0.4	3.0
83235 CC Mod Diff (ODO3920)	0.7	11.4
83236 CC NAT (ODO3920)	0.5	10.7
83237 CC Gr.2 ascend colon (ODO3921)	0.0	2.8
83238 CC NAT (ODO3921)	0.0	2.8
83241 CC from Partial Hepatectomy (ODO4309)	0.3	3.9
83242 Liver NAT (ODO4309)	0.0	0.3
87472 Colon mets to lung (OD04451-01)	0.4	7.5
87473 Lung NAT (OD04451-02)	0.2	4.3
Normal Prostate Clontech A+ 6546-1	1.7	0.0
84140 Prostate Cancer (OD04410)	1.3	10.8
84141 Prostate NAT (OD04410)	0.9	21.8
87073 Prostate Cancer (OD04720-01)	**100.0**	43.8
87074 Prostate NAT (OD04720-02)	0.9	19.8

Normal Lung GENPAK 061010	0.2	9.8
83239 Lung Met to Muscle (ODO4286)	0.0	0.0
83240 Muscle NAT (ODO4286)	0.6	5.4
84136 Lung Malignant Cancer (OD03126)	0.0	1.3
84137 Lung NAT (OD03126)	0.2	5.6
84871 Lung Cancer (OD04404)	0.0	0.8
84872 Lung NAT (OD04404)	0.6	5.0
84875 Lung Cancer (OD04565)	0.0	1.2
84876 Lung NAT (OD04565)	0.3	2.3
85950 Lung Cancer (OD04237-01)	0.4	6.0
85970 Lung NAT (OD04237-02)	0.0	4.9
83255 Ocular Mel Met to Liver (ODO4310)	0.0	1.4
83256 Liver NAT (ODO4310)	0.0	2.1
84139 Melanoma Mets to Lung (OD04321)	0.0	0.7
84138 Lung NAT (OD04321)	0.3	3.1
Normal Kidney GENPAK 061008	1.7	21.9
83786 Kidney Ca, Nuclear grade 2 (OD04338)	0.4	18.6
83787 Kidney NAT (OD04338)	0.6	10.5
83788 Kidney Ca Nuclear grade 1/2 (OD04339)	0.6	10.1
83789 Kidney NAT (OD04339)	1.1	16.8
83790 Kidney Ca, Clear cell type (OD04340)	0.4	6.2
83791 Kidney NAT (OD04340)	0.9	11.5
83792 Kidney Ca, Nuclear grade 3 (OD04348)	0.0	0.0
83793 Kidney NAT (OD04348)	0.4	8.7
87474 Kidney Cancer (OD04622-01)	0.0	0.6
87475 Kidney NAT (OD04622-03)	0.0	0.8
85973 Kidney Cancer (OD04450-01)	0.2	5.0
85974 Kidney NAT (OD04450-03)	0.3	6.1
Kidney Cancer Clontech 8120607	0.2	3.5
Kidney NAT Clontech 8120608	0.4	1.1
Kidney Cancer Clontech 8120613	0.2	2.8
Kidney NAT Clontech 8120614	0.1	5.4
Kidney Cancer Clontech 9010320	0.0	1.9
Kidney NAT Clontech 9010321	0.6	8.6
Normal Uterus GENPAK 061018	0.3	1.4
Uterus Cancer GENPAK 064011	1.1	17.0
Normal Thyroid Clontech A+ 6570-1	0.8	6.8
Thyroid Cancer GENPAK 064010	0.3	4.0
Thyroid Cancer INVITROGEN A302152	0.4	7.9
Thyroid NAT INVITROGEN A302153	0.3	9.0
Normal Breast GENPAK 061019	1.2	16.0
84877 Breast Cancer (OD04566)'	2.3	40.1
85975 Breast Cancer (OD04590-01)	1.2	17.8
85976 Breast Cancer Mets (OD04590-03)	1.2	12.3
87070 Breast Cancer Metastasis (OD04655-05)	1.7	23.2
GENPAK Breast Cancer 064006	0.8	15.8

Breast Cancer Res. Gen. 1024	7.5	**100.0**
Breast Cancer Clontech 9100266	0.8	7.1
Breast NAT Clontech 9100265	0.4	8.2
Breast Cancer INVITROGEN A209073	1.0	19.2
Breast NAT INVITROGEN A2090734	1.1	11.9
Normal Liver GENPAK 061009	0.0	3.8
Liver Cancer GENPAK 064003	0.2	1.2
Liver Cancer Research Genetics RNA 1025	0.0	3.7
Liver Cancer Research Genetics RNA 1026	0.0	1.4
Paired Liver Cancer Tissue Research Genetics RNA 6004-T	0.6	3.0
Paired Liver Tissue Research Genetics RNA 6004-N	0.1	0.6
Paired Liver Cancer Tissue Research Genetics RNA 6005-T	0.0	0.5
Paired Liver Tissue Research Genetics RNA 6005-N	0.0	0.3
Normal Bladder GENPAK 061001	0.2	7.7
Bladder Cancer Research Genetics RNA 1023	0.1	2.3
Bladder Cancer INVITROGEN A302173	0.2	3.0
87071 Bladder Cancer (OD04718-01)	0.0	1.3
87072 Bladder Normal Adjacent (OD04718-03)	0.9	19.1
Normal Ovary Res. Gen.	0.0	3.6
Ovarian Cancer GENPAK 064008	0.7	10.0
87492 Ovary Cancer (OD04768-07)	0.2	3.7
87493 Ovary NAT (OD04768-08)	0.2	1.9
Normal Stomach GENPAK 061017	1.2	15.4
Gastric Cancer Clontech 9060358	0.3	2.9
NAT Stomach Clontech 9060359	0.2	2.1
Gastric Cancer Clontech 9060395	0.4	8.2
NAT Stomach Clontech 9060394	0.3	4.2
Gastric Cancer Clontech 9060397	0.2	5.1
NAT Stomach Clontech 9060396	0.2	1.4
Gastric Cancer GENPAK 064005	0.2	6.8

Table 39. Panel 4.1D

Tissue Name	Relative Expression(%) 4.1x4tm6516f_ ag1539_a1	Tissue Name	Relative Expression(%) 4.1x4tm6516f_ ag1539_a1
93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.0	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.5	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting	0.9	93781_HUVEC	0.0

day 4-6 in IL-2	
93571_Secondary Tr1_resting day 4-6 in IL-2	0.6
93568_primary Th1_anti-CD28/anti-CD3	0.2
93569_primary Th2_anti-CD28/anti-CD3	0.7
93570_primary Tr1_anti-CD28/anti-CD3	0.0
93565_primary Th1_resting dy 4-6 in IL-2	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0
93567_primary Tr1_resting dy 4-6 in IL-2	1.1
93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3	1.4
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3	1.6
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.6
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	1.1
93354_CD4_none	2.9
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	1.3
93103_LAK cells_resting	1.5
93788_LAK cells_IL-2	1.6
93787_LAK cells_IL-2+IL-12	0.4
93789_LAK cells_IL-2+IFN gamma	2.1
93790_LAK cells_IL-2+ IL-18	2.0
93104_LAK cells_PMA/ionomycin and IL-18	0.2
93578_NK Cells IL-2_resting	0.4
93109_Mixed Lymphocyte Reaction_Two Way MLR	2.6
93110_Mixed Lymphocyte Reaction_Two Way MLR	2.2
(Endothelial)_IL-11	
93583_Lung Microvascular Endothelial Cells_none	0.7
93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.2
92662_Microvascular Dermal endothelium_none	0.3
92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	3.6
93347_Small Airway Epithelium_none	0.7
93348_Small Airway Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.9
92668_Coronery Artery SMC_resting	0.4
92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.7
93107_astrocytes_resting	6.1
93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	2.8
92666_KU-812 (Basophil)_resting	0.0
92667_KU-812 (Basophil)_PMA/ionoycin	0.0
93579_CCD1106 (Keratinocytes)_none	0.7
93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.7
93791_Liver Cirrhosis	0.8
93577_NCI-H292	5.3
93358_NCI-H292_IL-4	2.7
93360_NCI-H292_IL-9	5.6
93359_NCI-H292_IL-13	0.0
93357_NCI-H292_IFN gamma	0.8
93777_HPAEC_-	0.0
93778_HPAEC_IL-1 beta/TNA alpha	0.0

93111_Mixed Lymphocyte Reaction_Two Way MLR	0.4	93254_Normal Human Lung Fibroblast_none	8.5
93112_Mononuclear Cells (PBMCs)_resting	0.5	93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.3
93113_Mononuclear Cells (PBMCs)_PWM	0.0	93257_Normal Human Lung Fibroblast_IL-4	0.8
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0	93256_Normal Human Lung Fibroblast_IL-9	3.1
93249_Ramos (B cell)_none	0.0	93255_Normal Human Lung Fibroblast_IL-13	0.5
93250_Ramos (B cell)_ionomycin	0.0	93258_Normal Human Lung Fibroblast_IFN gamma	1.6
93349_B lymphocytes_PWM	0.0	93106_Dermal Fibroblasts CCD1070_resting	0.0
93350_B lymphoytes_CD40L and IL-4	1.1	93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	1.0
92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.8	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	1.3
93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93772_dermal fibroblast_IFN gamma	3.3
93356_Dendritic Cells_none	0.4	93771_dermal fibroblast_IL-4	2.7
93355_Dendritic Cells_LPS 100 ng/ml	0.0	93892_Dermal fibroblasts_none	4.1
93775_Dendritic Cells_anti-CD40	0.0	99202_Neutrophils_TNFa+LPS	0.4
93774_Monocytes_resting	1.3	99203_Neutrophils_none	1.2
93776_Monocytes_LPS 50 ng/ml	0.3	735010_Colon_normal	4.4
93581_Macrophages_resting	0.3	735019_Lung_none	5.6
93582_Macrophages_LPS 100 ng/ml	0.0	64028-1_Thymus_none	25.8
93098_HUVEC (Endothelial)_none	0.0	64030-1_Kidney_none	**100.0**
93099_HUVEC (Endothelial)_starved	0.0		

Table 40. Panel CNSD.01

Tissue Name	Relative Expression(%) cns_1x4tm654 8f_ag1539_a2	Tissue Name	Relative Expression(%) cns_1x4tm654 8f_ag1539_a2
102633_BA4 Control	29.0	102605_BA17 PSP	35.0
102641_BA4 Control2	39.6	102612_BA17 PSP2	17.3
102625_BA4 Alzheimer's2	19.1	102637_Sub Nigra Control	29.8
102649_BA4 Parkinson's	69.4	102645_Sub Nigra Control2	10.3
102656_BA4 Parkinson's2	62.4	102629_Sub Nigra Alzheimer's2	10.7

102664_BA4 Huntington's	21.0	102660_Sub Nigra Parkinson's2	26.1
102671_BA4 Huntington's2	8.5	102667_Sub Nigra Huntington's	65.0
102603_BA4 PSP	19.8	102674_Sub Nigra Huntington's2	11.5
102610_BA4 PSP2	18.2	102614_Sub Nigra PSP2	0.0
102588_BA4 Depression	27.0	102592_Sub Nigra Depression	7.0
102596_BA4 Depression2	17.5	102599_Sub Nigra Depression2	5.0
102634_BA7 Control	53.0	102636_Glob Palladus Control	19.8
102642_BA7 Control2	58.2	102644_Glob Palladus Control2	12.3
102626_BA7 Alzheimer's2	18.5	102620_Glob Palladus Alzheimer's	8.8
102650_BA7 Parkinson's	35.1	102628_Glob Palladus Alzheimer's2	49.1
102657_BA7 Parkinson's2	53.0	102652_Glob Palladus Parkinson's	89.9
102665_BA7 Huntington's	72.5	102659_Glob Palladus Parkinson's2	9.6
102672_BA7 Huntington's2	34.3	102606_Glob Palladus PSP	8.2
102604_BA7 PSP	70.3	102613_Glob Palladus PSP2	4.1
102611_BA7 PSP2	30.1	102591_Glob Palladus Depression	17.4
102589_BA7 Depression	14.3	102638_Temp Pole Control	7.1
102632_BA9 Control	34.9	102646_Temp Pole Control2	75.9
102640_BA9 Control2	73.9	102622_Temp Pole Alzheimer's	9.4
102617_BA9 Alzheimer's	15.5	102630_Temp Pole Alzheimer's2	17.1
102624_BA9 Alzheimer's2	19.8	102653_Temp Pole Parkinson's	38.3
102648_BA9 Parkinson's	58.0	102661_Temp Pole Parkinson's2	38.8
102655_BA9 Parkinson's2	66.2	102668_Temp Pole Huntington's	45.6
102663_BA9 Huntington's	52.5	102607_Temp Pole PSP	14.7
102670_BA9 Huntington's2	34.9	102615_Temp Pole PSP2	21.3
102602_BA9 PSP	21.1	102600_Temp Pole Depression2	9.0
102609_BA9 PSP2	6.9	102639_Cing Gyr Control	39.0
102587_BA9 Depression	20.9	102647_Cing Gyr Control2	48.6
102595_BA9 Depression2	9.6	102623_Cing Gyr Alzheimer's	12.4
102635_BA17 Control	74.2	102631_Cing Gyr Alzheimer's2	11.1
102643_BA17 Control2	**100.0**	102654_Cing Gyr Parkinson's	18.0
102627_BA17 Alzheimer's2	23.3	102662_Cing Gyr Parkinson's2	32.8
102651_BA17 Parkinson's	82.8	102669_Cing Gyr Huntington's	81.6
102658_BA17 Parkinson's2	91.3	102676_Cing Gyr Huntington's2	23.9
102666_BA17 Huntington's	59.8	102608_Cing Gyr PSP	19.6
102673_BA17 Huntington's2	36.6	102616_Cing Gyr PSP2	7.1
102590_BA17 Depression	31.9	102594_Cing Gyr Depression	19.1
102597_BA17 Depression2	46.3	102601_Cing Gyr Depression2	14.9

Panel 1.2 Summary: The 21629632_EXT gene shows rather ubiquitous expression across the samples on this panel, with highest expression in cerebral cortex (Ctmin=25) and hippocampus. See Panel 1.3D summary for explanation.

Panel 1.3D Summary: The expression of the 21629632_EXT gene is most highly represented in the samples of brain tissue and the sample of fetal muscle. The latter profile is of particular interest in that it differs significantly from that of the adult skeletal muscle. This difference implies that this protein may function to enhance muscular growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Thus, therapeutic modulation of this gene could be useful in treatment of muscular related disease. For instance treatment of weak or dystrophic muscle with the protein encoded by this gene could restore muscle mass or function. The 21629632_EXT transcript also shows highly preferential expression in brain, especially in the hippocampus and cerebral cortex where the expression is fairly high (CT = 29.5). The protein encoded by the 21629632_EXT gene appears to be a GPCR, making it an excellent small molecule target. Both the hippocampus and cerebral cortex are affected by neurodegeneration in Alzheimer's disease; thus this molecule is an excellent candidate for a drug target for the treatment/prevention of Alzheimer's disease, and may also be useful for memory enhancement/processing in healthy subjects.

Panel 2D Summary: The expression profile of the 21629632_EXT gene on this panel was assessed in duplicate runs, in which one run, designated as 2Dtm2349f was deemed to be erroneous. It appears that one sample of prostate cancer is contaminated with genomic DNA causing a skew in the data presentation. If this run is disregarded this gene appears to be expressed to a significant degree in a number of tissues. Particularly predominant is its expression in breast cancer and to a lesser degree in prostate cancer. Thus, therapeutic modulation of this gene may be of use in the treatment of breast cancer and/or prostate cancer or other breast and/or prostate related disease.

Panel 4.1D Summary: The 21629632_EXT gene is expressed at high levels in the kidney and at somewhat lower levels in the thymus. The 21629632_EXT transcript, the protein encoded for by the transcript, or antibodies designed with the protein could be used to identify kidney and thymus tissue.

Panel CNSD.01 Summary: An examination of 21629632_EXT gene expression in 8 brain regions across 12 individuals confirms that this protein is expressed in the brain of most, if not all individuals including those suffering from neurologic/psychiatric disease. Utility as a drug target would benefit from likely expression in most disease states.

I. GPCR10 (also known as 18234044_EXT)

Expression of gene 18234044_EXT was assessed using the primer-probe set Ag1539, described in Table 41. Results of the RTQ-PCR run are shown in Table 42.

Table 41. Probe name Ag1283

Primers	Sequences	TM	Length	Start Position	SEQ ID NO:
Forward	5'-TGATGGGACTCTTCAGACAATC-3'	59.1	22	69	104
Probe	FAM-5'-AACATCCAATGGCCAATATCACCTGG-3'-TAMRA	69.3	26	93	105
Reverse	5'-AAGAGTCCCAACAGGATGAAAT-3'	59	22	144	106

Table 42. Panel 4.1D

Tissue Name	Relative Expression(%) 4.1dx4tm6521f_ag1283_a1	Tissue Name	Relative Expression(%) 4.1dx4tm6521f_ag1283_a1
93768_Secondary Th1_anti-CD28/anti-CD3	0.0	93100_HUVEC (Endothelial)_IL-1b	0.0
93769_Secondary Th2_anti-CD28/anti-CD3	0.0	93779_HUVEC (Endothelial)_IFN gamma	0.0
93770_Secondary Tr1_anti-CD28/anti-CD3	0.0	93102_HUVEC (Endothelial)_TNF alpha + IFN gamma	0.0
93573_Secondary Th1_resting day 4-6 in IL-2	0.0	93101_HUVEC (Endothelial)_TNF alpha + IL4	0.0
93572_Secondary Th2_resting day 4-6 in IL-2	0.0	93781_HUVEC (Endothelial)_IL-11	0.0
93571_Secondary Tr1_resting day 4-6 in IL-2	0.0	93583_Lung Microvascular Endothelial Cells_none	0.0
93568_primary Th1_anti-CD28/anti-CD3	0.0	93584_Lung Microvascular Endothelial Cells_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93569_primary Th2_anti-CD28/anti-CD3	0.0	92662_Microvascular Dermal endothelium_none	0.0
93570_primary Tr1_anti-CD28/anti-CD3	0.0	92663_Microsvasular Dermal endothelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93565_primary Th1_resting dy 4-6 in IL-2	0.0	93773_Bronchial epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml) **	0.0
93566_primary Th2_resting dy 4-6 in IL-2	0.0	93347_Small Airway Epithelium_none	0.0
93567_primary Tr1_resting dy	0.0	93348_Small Airway	0.0

4-6 in IL-2	
93351_CD45RA CD4 lymphocyte_anti-CD28/anti-CD3	0.0
93352_CD45RO CD4 lymphocyte_anti-CD28/anti-CD3	0.0
93251_CD8 Lymphocytes_anti-CD28/anti-CD3	0.0
93353_chronic CD8 Lymphocytes 2ry_resting dy 4-6 in IL-2	0.0
93574_chronic CD8 Lymphocytes 2ry_activated CD3/CD28	0.0
93354_CD4_none	0.0
93252_Secondary Th1/Th2/Tr1_anti-CD95 CH11	0.0
93103_LAK cells_resting	0.0
93788_LAK cells_IL-2	0.0
93787_LAK cells_IL-2+IL-12	0.0
93789_LAK cells_IL-2+IFN gamma	0.0
93790_LAK cells_IL-2+ IL-18	0.0
93104_LAK cells_PMA/ionomycin and IL-18	0.0
93578_NK Cells IL-2_resting	0.0
93109_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93110_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93111_Mixed Lymphocyte Reaction_Two Way MLR	0.0
93112_Mononuclear Cells (PBMCs)_resting	0.0
93113_Mononuclear Cells (PBMCs)_PWM	0.0
93114_Mononuclear Cells (PBMCs)_PHA-L	0.0
93249_Ramos (B cell)_none	0.0
93250_Ramos (B cell)_ionomycin	0.0
93349_B lymphocytes_PWM	0.0
93350_B lymphoytes_CD40L and IL-4	0.0
Epithelium_TNFa (4 ng/ml) and IL1b (1 ng/ml)	
92668_Coronery Artery SMC_resting	0.0
92669_Coronery Artery SMC_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
93107_astrocytes_resting	0.0
93108_astrocytes_TNFa (4 ng/ml) and IL1b (1 ng/ml)	0.0
92666_KU-812 (Basophil)_resting	0.0
92667_KU-812 (Basophil)_PMA/ionoycin	0.0
93579_CCD1106 (Keratinocytes)_none	0.0
93580_CCD1106 (Keratinocytes)_TNFa and IFNg **	0.0
93791_Liver Cirrhosis	0.0
93577_NCI-H292	0.0
93358_NCI-H292_IL-4	0.0
93360_NCI-H292_IL-9	0.0
93359_NCI-H292_IL-13	0.0
93357_NCI-H292_IFN gamma	0.0
93777_HPAEC_-	0.0
93778_HPAEC_IL-1 beta/TNA alpha	0.0
93254_Normal Human Lung Fibroblast_none	0.0
93253_Normal Human Lung Fibroblast_TNFa (4 ng/ml) and IL-1b (1 ng/ml)	0.0
93257_Normal Human Lung Fibroblast_IL-4	0.0
93256_Normal Human Lung Fibroblast_IL-9	0.0
93255_Normal Human Lung Fibroblast_IL-13	0.0
93258_Normal Human Lung Fibroblast_IFN gamma	0.0
93106_Dermal Fibroblasts CCD1070_resting	0.0
93361_Dermal Fibroblasts CCD1070_TNF alpha 4 ng/ml	0.0

92665_EOL-1 (Eosinophil)_dbcAMP differentiated	0.0	93105_Dermal Fibroblasts CCD1070_IL-1 beta 1 ng/ml	0.0
93248_EOL-1 (Eosinophil)_dbcAMP/PMAionomycin	0.0	93772_dermal fibroblast_IFN gamma	0.9
93356_Dendritic Cells_none	0.0	93771_dermal fibroblast_IL-4	0.0
93355_Dendritic Cells_LPS 100 ng/ml	0.0	93892_Dermal fibroblasts_none	1.0
93775_Dendritic Cells_anti-CD40	0.0	99202_Neutrophils_TNFa+LPS	0.0
93774_Monocytes_resting	0.0	99203_Neutrophils_none	0.0
93776_Monocytes_LPS 50 ng/ml	0.0	735010_Colon_normal	0.0
93581_Macrophages_resting	0.0	735019_Lung_none	1.1
93582_Macrophages_LPS 100 ng/ml	0.0	64028-1_Thymus_none	8.4
93098_HUVEC (Endothelial)_none	0.0	64030-1_Kidney_none	**100.0**
93099_HUVEC (Endothelial)_starved	0.0		

Panel 2.2 Summary: Expression of the 18234044_EXT gene was low/undetectable (CT values > 35) in all samples on this panel and thus has not been shown.

Panel 4.1D Summary: The 18234044_EXT gene is expressed at high levels in the kidney. The 18234044_EXT transcript, the protein encoded for by the transcript or antibodies designed with the protein could be used to identify kidney tissue.

Example 3. SNP analysis of GPCRX clones

SeqCallingTM Technology: cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, cell lines, primary cells or tissue cultured primary cells and cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression for example, growth factors, chemokines, steroids. The cDNA thus derived was then sequenced using CuraGen's proprietary SeqCalling technology. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled with themselves and with public ESTs using bioinformatics programs to generate CuraGen's human SeqCalling database of SeqCalling assemblies. Each assembly contains one or more overlapping cDNA sequences derived from one or more human samples. Fragments and ESTs were included as components for an assembly when the extent of identity with another component of the

assembly was at least 95% over 50 bp. Each assembly can represent a gene and/or its variants such as splice forms and/or single nucleotide polymorphisms (SNPs) and their combinations.

Variant sequences are included in this application. A variant sequence can include a single nucleotide polymorphism (SNP). A SNP can, in some instances, be referred to as a "cSNP" to denote that the nucleotide sequence containing the SNP originates as a cDNA. A SNP can arise in several ways. For example, a SNP may be due to a substitution of one nucleotide for another at the polymorphic site. Such a substitution can be either a transition or a transversion. A SNP can also arise from a deletion of a nucleotide or an insertion of a nucleotide, relative to a reference allele. In this case, the polymorphic site is a site at which one allele bears a gap with respect to a particular nucleotide in another allele. SNPs occurring within genes may result in an alteration of the amino acid encoded by the gene at the position of the SNP. Intragenic SNPs may also be silent, however, in the case that a codon including a SNP encodes the same amino acid as a result of the redundancy of the genetic code. SNPs occurring outside the region of a gene, or in an intron within a gene, do not result in changes in any amino acid sequence of a protein but may result in altered regulation of the expression pattern for example, alteration in temporal expression, physiological response regulation, cell type expression regulation, intensity of expression, stability of transcribed message.

Method of novel SNP Identification: SNPs are identified by analyzing sequence assemblies using CuraGen's proprietary SNPTool algorithm. SNPTool identifies variation in assemblies with the following criteria: SNPs are not analyzed within 10 base pairs on both ends of an alignment; Window size (number of bases in a view) is 10; The allowed number of mismatches in a window is 2; Minimum SNP base quality (PHRED score) is 23; Minimum number of changes to score an SNP is 2/assembly position. SNPTool analyzes the assembly and displays SNP positions, associated individual variant sequences in the assembly, the depth of the assembly at that given position, the putative assembly allele frequency, and the SNP sequence variation. Sequence traces are then selected and brought into view for manual validation. The consensus assembly sequence is imported into CuraTools along with variant sequence changes to identify potential amino acid changes resulting from the SNP sequence variation. Comprehensive SNP data analysis is then exported into the SNPCalling database.

Method of novel SNP Confirmation: SNPs are confirmed employing a validated method know as Pyrosequencing (Pyrosequencing, Westborough, MA). Detailed protocols for Pyrosequencing can be found in: Alderborn et al. Determination of Single Nucleotide Polymorphisms by Real-time Pyrophosphate DNA Sequencing. (2000). *Genome Research*. 10, Issue 8, August. 1249-1265. In brief, Pyrosequencing is a real time primer extension process

of genotyping. This protocol takes double-stranded, biotinylated PCR products from genomic DNA samples and binds them to streptavidin beads. These beads are then denatured producing single stranded bound DNA. SNPs are characterized utilizing a technique based on an indirect bioluminometric assay of pyrophosphate (PPi) that is released from each dNTP upon DNA chain elongation. Following Klenow polymerase-mediated base incorporation, PPi is released and used as a substrate, together with adenosine 5'-phosphosulfate (APS), for ATP sulfurylase, which results in the formation of ATP. Subsequently, the ATP accomplishes the conversion of luciferin to its oxi-derivative by the action of luciferase. The ensuing light output becomes proportional to the number of added bases, up to about four bases. To allow processivity of the method dNTP excess is degraded by apyrase, which is also present in the starting reaction mixture, so that only dNTPs are added to the template during the sequencing. The process has been fully automated and adapted to a 96-well format, which allows rapid screening of large SNP panels.

The DNA and protein sequences for the novel single nucleotide polymorphic variants are reported. Variants are reported individually but any combination of all or a select subset of variants are also included. In addition, the positions of the variant bases and the variant amino acid residues are underlined.

GPCR2 SNP

The nucleotide sequence of variant 13374652 (bold and underlined) has a T/G single nucleotide polymorphism ("SNP) as shown in Table 43. The SNP occurrence in nucleotide position 381 T->G results in a protein sequence variant in amino acid position 127 resulting in Cys to Trp.

Table 43 Variant of GPCR2 nucleotide sequence.

```
   1 ATGACCATGGAAAATTATTCTATGGCAGCTCAGTTTGTCTTAGATGGTTTAACACAGCAAGCAGAGCTCCAGCTGCCCCT
  81 CTTCCTCCTGTTCCTGGGAATCTATGTGGTCACAGTAGTGGGCAACCTGGGCATGATTCTCCTGATTGCAGTCAGCCCTC
 161 TACTTCACACCCCCATGTACTATTTCCTCAGCAGCTTGTCCTTCGTCGATTTCTGCTATTCCTCTGTCATTACTCCCAAA
 241 ATGCTGGTGAACTTCCTAGGAAAGAAGAATACAATCCTTTACTCTGAGTGCATGGTCCAGCTCTTTTTCTTTGTGGTCTT
 321 TGTGGTGGCTGAGGGTTACCTCCTGACTGCCATGGCATATGATCGCTATGTTGCCATCTGGAGCCCACTGCTTTATAATG
 401 CGATCATGTCCTCATGGGTCTGCTCACTGCTAGTGCTGGCTGCCTTCTTCTTGGGCTTTCTCTCTGCCTTGACTCATACA
 481 AGTGCCATGATGAAACTGTCCTTTTGCAAATCCCACATTATCAACCATTACTTCTGTGATGTTCTTCCCCTCCTCAATCT
 561 CTCCTGCTCCAACACACACCTCAATGAGCTTCTACTTTTTATCATTGCGGGGTTTAACACCTTGGTGCCCACCCTAGCTG
 641 TTGCTGTCTCCTATGCCTTCATCCTCTACAGCATCCTTCACATCCGCTCCTCAGAGGGCCGGTCCAAAGCTTTTGGAACA
 721 TGCAGCTCTCATCTCATGGCTGTGGTGATCTTCTTTGGGTCCATTACCTTCATGTATTTCAAGCCCCCTTCAAGTAACTC
 801 CCTGGACCAGGAGAAGGTGTCCTCTGTGTTCTACACCACGGTGATCCCCATGCTGAACCCTTTAATATACAGTCTGAGGA
 881 ATAAGGATGTGAAGAAAGCATTAAGGAAGGTCTTAGTAGGAAAATGA
```

GPCR4 SNP

The nucleotide sequence of SNP variant 13374653 (bold and underlined) has a C/T as shown in Table 44. The SNP occurrence in nucleotide position 348 C->T does not result in a protein sequence variant.

Table 44 Variant of GPCR4 nucleotide sequence.

6A1. Nucleotide sequence of variant 13374653 (underlined).

```
  1 ATGCTGGCTAGAAACAACTCCTTAGTGACTGAATTTATTCTTGCTGGATTAACAGATCATCCAGAGTTCCAGCAACCCCT
 81 CTTTTTCCTGTTTCTAGTGGTCTACATTGTCACCATGGTAGGCAACCTTGGCTTGATCATTCTTTTCGGTCTAAATTCTC
161 ACCTCCACACACCAATGTACTATTTCCTCTTCAATCTCTCCTTCATTGATCTCTGTTACTCCTCTGTTTTCACTCCCAAA
241 ATGCTAATGAACTTTGTATCAAAAAAGAATATTATCTCCTATGTTGGGTGCATGACTCAGCTGTTTTTCTTTCTCTTTTT
321 TGTCATCTCTGAATGTTACATGTTGACTTCAATGGCATATGATCGCTATGTGGCCATCTGTAATCCATTGCTGTATAAGG
401 TCACCATGTCCCATCAGGTCTGTTCTATGCTCACTTTTGCTGCTTACATAATGGGATTGGCTGGAGCCACGGCCCACACC
481 GGGTGCATGCTTAGACTCACCTTCTGCAGTGCTAATATCATCAACCATTACTTGTGTGACATACTCCCCCTCCTCCAGCT
561 TTCCTGCACCAGCACCTATGTCAACGAGGTGGTTGTTCTCATTGTTGTGGGTATTAATATCATGGTACCCAGTTGTACCA
641 TCCTCATTTCTTATGTTTTCATTGTCACTAGCATTCTTCATATCAAATCCACTCAAGGAAGATCAAAAGCCTTCAGTACT
721 TGTAGCTCTCATGTCATTGCTCTGTCTCTGTTTTTTGGGTCAGCGGCATTCATGTATATTAAATATTCTTCTGGATCTAT
801 GGAGCAGGGAAAAGTTTCTTCTGTTTTCTACACTAATGTGGTGCCCATGCTCAATCCTCTCATCTACAGTTTGAGGAACA
881 AGGATGTCAAAGTTGCACTGAGGAAAGCTCTGATTAAAATTCAGAGAAGAAATATATTCTAA
```

EQUIVALENTS

Although particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims, which follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made to the invention without departing from the spirit and scope of the invention as defined by the claims. The choice of nucleic acid starting material, clone of interest, or library type is believed to be a matter of routine for a person of ordinary skill in the art with knowledge of the embodiments described herein. Other aspects, advantages, and modifications considered to be within the scope of the following claims.

WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

(b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;

(c) an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28; and

(d) a variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence.

2 The polypeptide of claim 1, wherein said polypeptide comprises the amino acid sequence of a naturally-occurring allelic variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28.

3. The polypeptide of claim 2, wherein said allelic variant comprises an amino acid sequence that is the translation of a nucleic acid sequence differing by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27.

4. The polypeptide of claim 1, wherein the amino acid sequence of said variant comprises a conservative amino acid substitution.

5. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

(b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;

(c) an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

(d) a variant of an amino acid sequence selected from the group consisting SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence;

(e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising an amino acid sequence chosen from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, or a variant of said polypeptide, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence; and

(f) a nucleic acid molecule comprising the complement of (a), (b), (c), (d) or (e).

6. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises the nucleotide sequence of a naturally-occurring allelic nucleic acid variant.

7. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule encodes a polypeptide comprising the amino acid sequence of a naturally-occurring polypeptide variant.

8. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27.

9. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27;

(b) a nucleotide sequence differing by one or more nucleotides from a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, provided that no more than 20% of the nucleotides differ from said nucleotide sequence;

(c) a nucleic acid fragment of (a); and

(d) a nucleic acid fragment of (b).

10. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule hybridizes under stringent conditions to a nucleotide sequence chosen from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 20, 22, 24, 26 and 27, or a complement of said nucleotide sequence.

11. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence comprising a coding sequence differing by one or more nucleotide sequences from a coding sequence encoding said amino acid sequence, provided that no more than 20% of the nucleotides in the coding sequence in said first nucleotide sequence differ from said coding sequence;

(b) an isolated second polynucleotide that is a complement of the first polynucleotide; and

(c) a nucleic acid fragment of (a) or (b).

12. A vector comprising the nucleic acid molecule of claim 11.

13. The vector of claim 12, further comprising a promoter operably-linked to said nucleic acid molecule.

14. A cell comprising the vector of claim 12.

15. An antibody that binds immunospecifically to the polypeptide of claim 1.

16. The antibody of claim 15, wherein said antibody is a monoclonal antibody.

17. The antibody of claim 15, wherein the antibody is a humanized antibody.

18. A method for determining the presence or amount of the polypeptide of claim 1 in a sample, the method comprising:

(a) providing the sample;

(b) contacting the sample with an antibody that binds immunospecifically to the polypeptide; and

(c) determining the presence or amount of antibody bound to said polypeptide, thereby determining the presence or amount of polypeptide in said sample.

19. A method for determining the presence or amount of the nucleic acid molecule of claim 5 in a sample, the method comprising:

(a) providing the sample;

(b) contacting the sample with a probe that binds to said nucleic acid molecule; and

(c) determining the presence or amount of the probe bound to said nucleic acid molecule, thereby determining the presence or amount of the nucleic acid molecule in said sample.

20. The method of claim 19 wherein presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type.

21. The method of claim 20 wherein the cell or tissue type is cancerous.

22. A method of identifying an agent that binds to a polypeptide of claim 1, the method comprising:

(a) contacting said polypeptide with said agent; and

(b) determining whether said agent binds to said polypeptide.

23. The method of claim 22 wherein the agent is a cellular receptor or a downstream effector.

24. A method for identifying an agent that modulates the expression or activity of the polypeptide of claim 1, the method comprising:

(a) providing a cell expressing said polypeptide;

(b) contacting the cell with said agent, and

(c) determining whether the agent modulates expression or activity of said polypeptide,

whereby an alteration in expression or activity of said peptide indicates said agent modulates expression or activity of said polypeptide.

25. A method for modulating the activity of the polypeptide of claim 1, the method comprising contacting a cell sample expressing the polypeptide of said claim with a compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.

26. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the polypeptide of claim 1 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

27. The method of claim 26 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

28. The method of claim 26 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

29. The method of claim 26, wherein said subject is a human.

30. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the nucleic acid of claim 5 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

31. The method of claim 30 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

32. The method of claim 30 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

33. The method of claim 30, wherein said subject is a human.

34. A method of treating or preventing a GPCRX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the antibody of claim 15 in an amount sufficient to treat or prevent said GPCRX-associated disorder in said subject.

35. The method of claim 34 wherein the disorder is diabetes.

36. The method of claim 34 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

37. The method of claim 34, wherein the subject is a human.

38. A pharmaceutical composition comprising the polypeptide of claim 1 and a pharmaceutically-acceptable carrier.

39. A pharmaceutical composition comprising the nucleic acid molecule of claim 5 and a pharmaceutically-acceptable carrier.

40. A pharmaceutical composition comprising the antibody of claim 15 and a pharmaceutically-acceptable carrier.

41. A kit comprising in one or more containers, the pharmaceutical composition of claim 38.

42. A kit comprising in one or more containers, the pharmaceutical composition of claim 39.

43. A kit comprising in one or more containers, the pharmaceutical composition of claim 40.

44. A method for determining the presence of or predisposition to a disease associated with altered levels of the polypeptide of claim 1 in a first mammalian subject, the method comprising:

(a) measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and

(b) comparing the amount of said polypeptide in the sample of step (a) to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, said disease;

wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to said disease.

45. The method of claim 44 wherein the predisposition is to cancers.

46. A method for determining the presence of or predisposition to a disease associated with altered levels of the nucleic acid molecule of claim 5 in a first mammalian subject, the method comprising:

(a) measuring the amount of the nucleic acid in a sample from the first mammalian subject; and

(b) comparing the amount of said nucleic acid in the sample of step (a) to the amount of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease;

wherein an alteration in the level of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

47. The method of claim 46 wherein the predisposition is to a cancer.

48. A method of treating a pathological state in a mammal, the method comprising administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide comprising an amino acid sequence of at least one of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, or a biologically active fragment thereof.

49. A method of treating a pathological state in a mammal, the method comprising administering to the mammal the antibody of claim 15 in an amount sufficient to alleviate the pathological state.

50. A method for the screening of a candidate substance interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, or fragments or variants thereof, comprises the following steps:

a) providing a polypeptide selected from the group consisting of the sequences of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, or a peptide fragment or a variant thereof;

b) obtaining a candidate substance;

c) bringing into contact said polypeptide with said candidate substance; and

d) detecting the complexes formed between said polypeptide and said candidate substance.

51. A method for the screening of ligand molecules interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein said method comprises:

a) providing a recombinant eukaryotic host cell containing a nucleic acid encoding a polypeptide selected from the group consisting of the polypeptides comprising the amino acid sequences SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

b) preparing membrane extracts of said recombinant eukaryotic host cell;

c) bringing into contact the membrane extracts prepared at step b) with a selected ligand molecule; and

d) detecting the production level of second messengers metabolites.

52. A method for the screening of ligand molecules interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein said method comprises:

a) providing an adenovirus containing a nucleic acid encoding a polypeptide selected from the group consisting of polypeptides comprising the amino acid sequences SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

b) infecting an olfactory epithelium with said adenovirus;

c) bringing into contact the olfactory epithelium b) with a selected ligand molecule; and

d) detecting the increase of the response to said ligand molecule.

52. A method for the screening of ligand molecules interacting with an olfactory receptor polypeptide selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28, wherein said method comprises:

a) providing an adenovirus containing a nucleic acid encoding a polypeptide selected from the group consisting of polypeptides comprising the amino acid sequences SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 21, 23, 25 and 28;

b) infecting an olfactory epithelium with said adenovirus;

c) bringing into contact the olfactory epithelium b) with a selected ligand molecule; and

d) detecting the increase of the response to said ligand molecule.